

# BESZÉDKUTATÁS 2018

26.

BUDAPEST

# **BESZÉDKUTATÁS 2018**

**26.**

**Szerkesztette:  
Gósy Mária**

MTA Nyelvtudományi Intézet  
Budapest, 2018

A szerkesztőbizottság elnöke: Gósy Mária (MTA Nyelvtudományi Intézet)  
A szerkesztőbizottság tagjai: Bunta Ferenc (University of Houston)  
Horváth Viktória (MTA Nyelvtudományi Intézet)  
Huntley Bahr, Ruth (University of South Florida)  
Mády Katalin (MTA Nyelvtudományi Intézet)  
Markó Alexandra (Eötvös Loránd Tudományegyetem)  
Olaszy Gábor (Budapesti Műszaki és Gazdaságtudományi Egyetem)  
Siptár Péter (Eötvös Loránd Tudományegyetem)  
Trouvain, Jürgen (University of Saarland)  
Vago, Robert (Queens College és City University of New York)

A korpusz szerkesztésében részt vettek: Bóna Judit  
Gyarmathy Dorottya  
Lakatos Boglárka

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztés: Krepsz Valéria

A kötet kiadását az MTA Könyv- és Folyóirat-kiadó Bizottsága támogatta.

A folyóiratot az MTMT indexeli és a REAL archiválja.

DOI: 10.15775/Beszkut.2018

ISSN 1218 - 8727

© MTA Nyelvtudományi Intézet

Felelős kiadó: Prószéky Gábor igazgató

---

## KÖSZÖNTÉS

### Olaszy Gábor 75 éves

A magyar fonetika tudományának egyik legjelentősebb, meghatározó személyisége Olaszy Gábor.

Több mint 30 éven át volt az MTA Nyelvtudományi Intézetében a Fonetikai Osztály tudományos munkatársa, illetve osztályvezetője. A beszéd elkötelezett kutatója. Munkássága a beszédtudomány számos területét felöleli. A magyar beszédre vonatkozóan éppen megkezdődött akusztikai-fonetikai kutatásokat elmélyítette, átfogó vizsgálatait, jelentős eredményei nagymértékben hozzájárultak a szaktudomány fejlődéséhez, illetve ahhoz, hogy a magyar fonetika nemzetközileg is elismertté vált.

Nem célunk minden eredményének a bemutatása, csupán néhány tényrt ismertetünk. Tudományos közleményeinek száma meghaladja a 170-et, ebből 58 idegen nyelven jelent meg. Négy könyvet írt, és több mint félezer ismert idézésről van tudomásunk. Számos szabadalmát fogadták el. Jól ismert előadója mind a hazai, mind a nemzetközi konferenciáknak. Tudományos tevékenysége a mai napig megszakíthatatlan. Folyamatosan kutat és publikál, részt vesz a minősítési bizottságok munkájában, bírálataival segíti a kéziratok szerzőit. Professzor emeritusként oktatja a beszédtudománnyal kapcsolatos tantárgyakat.

Nevéhez fűződik Kempelen Farkas beszélőgépének reprodukálása (munkatársával), amely nemzetközi szinten is nagy elismerést aratott. A magyar beszéd szintézisének, azaz a mesterséges beszéd előállításának terén végzett munkája egyedülálló; eredményei a tudományban mérföldkönek tekinthetők, a gyakorlatban pedig mind a mai napig hasznosulnak.

Olaszy Gábor az elmúlt évtizedekben mindig aktív résztvevője volt az általunk szervezett konferenciáknak, a folyóirat-indításoknak, kötet szerkesztéseknek, a magyar fonetika valamennyi jelentős eseményének. Ma is az, mindig segítőkész, megbízható, kitűnő kolléga, akinek munkájára, tanácsaira számíthatunk, és a jövőben is számítani fogunk. Köszönjük.

Születésnapja alkalmából minden fonetikus, beszédkutató és kolléga nevében kívánok sok (szakmai) örömet, további fonetikai kutatásokat és egészséget nagy szeretettel!

G. M.

## TARTALOM

Deme Andrea – Bartók Márton – Gráci Tekla Etelka – Markó Alexandra – Csapó Tamás Gábor: A magyar /h/ zöngésedése magánhangzók között .....	7
Bartók Márton: A gégeműködés variabilitása az érzelemkifejezés függvényében .....	30
Gocsál Ákos: Egyes testalkati paraméterek és a kitöltött szünetek formánsszerkezete közötti összefüggés vizsgálata férfi beszélőknél .....	63
Markó Alexandra – Bartók Márton – Gráci Tekla Etelka – Deme Andrea – Csapó Tamás Gábor: Mondathangsúlyos és hangsúlytalan helyzetű magánhangzók néhány artikulációs és akusztikai jellemzője a magyarban .....	85
Gósy Mária: Hiátusos kontextusok temporális mintázata gyermekek beszédében .....	110
Gyarmathy Dorottya – Horváth Viktória: A néma szünetek sajátosságai óvodások és kisiskolások spontán beszédében .....	134
Krepsz Valéria – Gósy Mária: A dichotikus szóészlelés életkorspe-cifikus jellemzői .....	156
Gósy Mária – Gyarmathy Dorottya: A beszédindulás késésének hatása a beszédfeldolgozásra .....	186
Bóna Judit: Disfluencies and disfluency clusters in cluttered, stuttered and typical speech .....	221
Zainkó Csaba – Gyires-Tóth Bálint – Németh Géza – Olaszy Gábor: Kísérletek a WaveNet módszer alkalmazására magyar beszéd- szintézishez .....	236

“Nyelvbtlás”-korpusz.....	250
Téves szó.....	251
Grammatikai hiba .....	255
Kontamináció .....	257
Téves kezdés .....	258
“Nyelvem hegyén van” jelenség .....	259
Perszeveráció .....	261
Anticipáció .....	262
Metatézis .....	263
Egyszerű nyelvbtlás .....	264
Több típusba sorolható jelenségek .....	265
Könyvismertető.....	266
Bóna Judit: Női beszéd – férfi beszéd a fonetikai és pszicholingvisztikai vizsgálatok tükrében • <i>Huszár Anna</i> .....	266
Szántó Anna: Beszédfeldolgozási folyamatok változásai a fejlesztés függvényében • <i>Désfalvi Ildikó</i> .....	268
Bóna Judit (szerk.) Új utak a gyermeknyelvi kutatásokban • <i>Hantó     Réka</i> .....	271
Gyarmathy Dorottya: Megakadásjelenségek a magyar spontán beszédben • <i>Bóna Judit</i> .....	274
Gósy Mária – Krepsz Valéria: Morfémák időzítési mintázatai a beszédben • <i>Horváth Viktória</i> .....	277



## A MAGYAR /h/ ZÖNGÉSEDÉSE MAGÁNHANGZÓK KÖZÖTT

**Deme Andrea<sup>1,4</sup> – Bartók Márton<sup>1,4</sup> – Gráci Tekla Etelka<sup>2,4</sup> –  
Markó Alexandra<sup>1,4</sup> – Csapó Tamás Gábor<sup>3,4</sup>**

<sup>1</sup> ELTE BTK Fonetikai Tanszék, <sup>2</sup> MTA Nyelvtudományi Intézet, <sup>3</sup> BME  
TMIT, <sup>4</sup> MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport

### Bevezetés

A magyar mássalhangzórendszerben egy kivételével az összes zöngétlen obstruens rendelkezik (kontrasztív) zöngés párral. Ez az egyetlen kivételes obstruens a zöngétlen laringális réshang, a /h/. Ugyanakkor, bár a /h/ fonológiai valóban nem áll oppozícióban zöngés laringális frikatívával, zöngés párja – a [ɦ] – bizonyos körülmények között, allofonikus variáció következtében megjelenik a felszínen (Siptár 2001; Siptár–Törkenczy 2000/2007). Habár a legújabb fonológiai megközelítések (vö. Siptár–Törkenczy 2000/2007) a fonémakészletben a zöngétlen veláris frikatívát /x/ definiálják a szóban forgó hangtípus reprezentánsaként, mivel a fonetikai irodalom egyöntetűen laringális frikativaként határozza meg (vö. Bolla 1995; Gósy 2004; Kassai 2005), a továbbiakban mi magunk is ez utóbbi megoldást követjük.

A jelen tanulmányban a /h/ zöngésségi allofonikus alternációját elemezzük fonetikai szempontból, azzal a céllal, hogy feltérképezzük a zöngés [ɦ] variáns előfordulásának fonetikai és fonológiai természetű előidézőit a magyarban. Ez a vizsgálat egy hosszabb kísérletsorozat része, melyben a fonológia és fonetika olyan széles körben elfogadni látszott állításait teszteljük, amelyekre a szakirodalom alaposabb szemügyre vétele után nem találunk megbízhatónak ítéltető empirikus bizonyítékokat. Az itt bemutatásra kerülő első kísérletben jól kontrollált laboratóriumi beszédanyagon három, a szakirodalomban feltételezett tényezőt vizsgálunk meg a /h/-realizáció zöngés részének arányára gyakorolt hatásuk tekintetében: két magánhangzós tulajdonságot, a nyíltságot (vagy függőleges nyelvhelyzetet vagy nyelvállásfokot) és az előlétséget (a nyelv vízszintes helyzetét), valamint a mondathangsúly szerepét. A kísérlet másodlagos célja képet adni a [ɦ] különleges zöngésségének, a leheletes vagy mormolt zöngének a fonetikai sajátosságairól. Ebből a célból egy további akusztikus paramétert is elemzünk, melyet hagyományosan a zöngeminőség jellemzésére (is) alkalmazunk: a harmonikus-zaj viszony (harmonics-to-noise ratio) paramétert.

Deme Andrea – Bartók Márton – Gráci Tekla Etelka – Markó Alexandra –  
Csapó Tamás Gábor 2018. A magyar /h/ zöngésedése magánhangzók között.  
*Beszédkutató* 2018. 7–29.



A magyar fonológiai szakirodalom a /h/ viselkedését főként a zöngességi hasonulásban való részvételével kapcsolatosan tárgyalja. A /h/ ugyanis ebben a szabályszerűségben aszimmetrikusan vesz részt: bár zöngétleníti a tőle balra álló zöngés obstruent, maga a /h/ egy tőle jobbra álló zöngés obstruens hatására nem zöngésedik (lásd Siptár 2001; Siptár–Törkenczy 2000/2007). A /h/ allofonikusnak tekintett zöngességi alternációját ugyanakkor alig részletezi a szakirodalom, ezt ugyanis szinte kizárólagosan fonetikai természetű és motivációjú folyamatok eredményének tartják (lásd pl. Szigetvári 1998) a következők szerint. A /h/ [h̥]-ként valósul meg (szinte) minden intervokális pozícióban (Siptár 2001; Siptár–Törkenczy 2000/2007) (erre a pozícióra a továbbiakban röviden *VhV*-ként hivatkozunk), valamint szonoráns és magánhangzó között is (Siptár 1994) (erre a pozícióra a továbbiakban röviden *szonhV*-ként hivatkozunk). A pusztán fonetikainak tekintett variációval kapcsolatban csak egy helyütt olvashatunk arról, hogy abban fonológiai faktorok is szerepet játszanának. Siptár (1994/2016: 213, 265) szerint ugyanis az intervokálisan realizálódó /h/ a lassú/gondozott beszédben kizárólag akkor zöngésedhet, ha a /h/-t tartalmazó szótag hangsúlytalan (pl. *tehén* [tɛɦe:n]), sőt ebben az esetben a zöngésedett /h/ egyúttal törlődhet is (pl. [tɛe:n]). Ha azonban a /h/ hangsúlyos szótagban realizálódik, mint például az *a hír* [pʰɦi:r] esetében, az említett beszédstílusban a /h/ nem zöngésedhet, és nem is törlődhet a hangsorból (\*[pʰi:r]). Ezzel szemben Siptár (1994/2016: 213, 265) ugyanakkor azt is megfogalmazza, hogy a /h/ gyors/lezser beszédben még az *a hír* [pʰɦi:r] típusú, hangsúlyos kontextusokban is zöngésedhet, ugyanis itt a szabály általánosodik, azaz nem veszi figyelembe a hangsúlyt.

A fentebb összefoglalt elméleti leírásokban olyan feltételezéseket találunk, melyek alapot kínálnak az empirikus tudományoknak, hogy az állításokban foglaltakat kísérletes úton is ellenőrizzék. Ezeket a feltételezéseket a következőképpen foglalhatjuk össze: 1. a /h/ zöngésedik intervokális kontextusban; 2. a /h/ intervokális zöngésedését befolyásolja a tartalmazó szótag hangsúlyossága; 3. a /h/ zöngésedik szonoráns és magánhangzó között. A /h/ zöngésedését befolyásoló itt említett hatások közül azonban – amint azt az alábbiakban látni fogjuk – valójában egyet sem támasztott még alá eddig jól kontrollált beszédanyagokon végzett kísérlet. Ráadásul bizonyos tényezők, például a *szonhV* kontextus hatását empirikus munkákban még egyáltalán nem elemezték korábban.

A fonetikai szakirodalomban a /h/-t illető kérdésekben meglehetősen régre datálható diskurzust találunk. Az egyik és talán legrégebbi kérdés ezek között a /h/ státuszát érinti, azt, hogy a /h/-t mássalhangzónak vagy magánhangzónak kell-e tekintenünk. Erről a dilemmáról átfogóan Lazicziusnál (1937, 1963/1979) olvashatunk. Idézett munkáiban Laziczius elsőként az 1920-as évekből (és az azelőtről) származó spekulációkat foglalja össze, melyek szerint szájüregi akadály (amelyet egy mássalhangzó alapvető jellemzőjének tartunk) híján, a glottális/laringális képzési helyű /h/ nem tekinthető mássalhangzónak, csak zöngétlen magánhangzónak (egészen pontosan az őt követő

zöngés magánhangzó zöngétlen párjának). Eztán a szerző arra is kitér, hogy ezt a nézetet később felülírta a tudományos közvélemény azzal a gyakorlatilag egyöntetű és a nemzetközi diskurzusban is teret nyert vélekedéssel, mely szerint képzési helye ellenére a /h/ mássalhangzóként való jellemzése a legjobb elemzési megoldás (a diskurzus korabeli, illetve azt követő évekre datálható nemzetközi vetületéhez lásd pl. Maddieson 1984; IPA 1999).

Egy másik pont, amelyben még a legfrissebb fonetikai kézikönyvek is meglehetősen változatos képet mutatnak, a /h/ feltételezett allofónjainak, még pontosabban az allofónok képzési helyének, valamint az allofónokat elicitáló fonetikai kontextusoknak a kérdése. Bolla (1995) egy férfi és egy női beszélő ejtéséről készített röntgenfelvételek alapján feltételez egy palatoveláris zöngétlen frikatívát, amelyet az IPA jelöléstechnikájától (és logikájától) merőben eltérően [ch]-val jelöl (holott általában az IPA jeleit alkalmazza), például az *ihlet* [ichlet]<sup>1</sup> és a *jacht* [jɔcht] szavakban. Bolla (1995) felvesz továbbá egy veláris zöngétlen frikatívát [x] például a *potrohos* [potroxɔ] szóban; egy faringális(!)<sup>2</sup> zöngétlen frikatívát [h] a *hát* [ha:t] és a *nátha* [na:thɔ] szavakban; valamint egy faringális(!) zöngés frikatívát [ɦ] a *nahát* [naɦa:t] szóban.

Gósy (2004) leírásában a következőképpen sorjázza a /h/ allofónjait. Feltételez egy veláris zöngétlen frikatívát [x] az *ihlet* [ixlet] szóban; egy hátrébb képzett „posztvelárist” (amelyet előbb [x]-vel (78), máshol – pontosabban – [χ]-vel (132) jelöl) veláris magánhangzók után, például a *potroh* [potrox] és a *doh* [dox] szavakban; egy veláris zöngés frikatívát [ɣ] a *dohos* [doɣɔ] szóban; egy laringális zöngétlen frikatívát [h] például a *hó* [ho:] szóban; valamint egy laringális zöngés frikatívát az *éhes* [e:ɦɛʃ] szóban.

Végezetül Kassai (2005) az eddigiektől részben eltérő módon ír egy néhol palatovelárisként [x̣] (104), máshol pedig palatálisként [ç] (105) meghatározott zöngétlen frikatíváról olyan kontextusokban, melyekben a /h/-t palatális magánhangzó előzi meg (pl. *ihlet* [ix̣let]/[içlet]). Felvesz továbbá egy veláris zöngétlen frikatívát [x] kóda pozícióban, veláris magánhangzó után (pl. *potroh* [potrox]); egy veláris zöngétlen frikatívát [ɣ] kóda pozícióban, zöngés mássalhangzók után (pl. *almanachban* [ɔlmɒnɒɰɰɰbɒn]); valamint egy laringális zöngétlen frikatívát [h] szó eleji és intervokális helyzetben. Ez

<sup>1</sup> Megjegyzendő, hogy az idézett példák transzkripciója nem feltétlenül egyezik meg teljes egészében az idézett szerzők jelöléstechnikájával, pusztán a szóban forgó frikatíva jelölésének bemutatásában követtük az idézett szerzők gyakorlatát.

<sup>2</sup> Habár Bolla (1995) a [h] szimbólumot a laringális/glottális zöngétlen réshang jeleként ismerteti, amelyet ennek megfelelően megkülönböztet a zöngétlen faringális réshang [h] jelétől (1995: 39), a magyar /h/ leírásakor a [h] szimbólumot alkalmazza, és következetesen faringális réshangként hivatkozik rá (1995: 37, 166, 212). Hasonlóképpen jár el a [ɦ] szimbólum alkalmazásakor is, melyet bár szintén laringálisként vezet be, faringálisként hivatkozik rá (1995: 37, 166, 212).

utóbbiról továbbá azt is megjegyzi, hogy „nemritkán mormolt zöngé színez[i] [fi]-vá” (Kassai 2005: 104), majd később így ír: „a zöngétlen gégerészhang, a /h/ két magánhangzó között vagy zengőhang szomszédságában maga is zöngésen hangzik” (151), ahol már kivétel nélküli mintázat meglétét sugallja. Zárasképp megjegyzendő, hogy Bollával szemben és Gósyhoz hasonlóan Kassai (2005) sem jelöli meg a feltételezett allofónokra utaló adatok empirikus forrását (ami összefügghet azzal, hogy az utóbbi két esetben tankönyvi szövegekről van szó).

Bár a /h/-allofónok képzési helyét érintő dilemmák megoldása messze túlmutat a jelen tanulmány keretein, a kérdés áttekintése mégis haszonnal jár a jelen kutatási kérdések szempontjából is. Feltűnő ugyanis, hogy az egyes szerzők által tételezett allofónok valójában nem csak a képzési helyük és az azt facilitáló fonetikai kontextus, de a zöngességük tekintetében sem mutatnak teljesen egységes képet. Ezzel elérkeztünk összefoglalónk utolsó és vizsgálatunk szempontjából legfontosabb pontjához, a /h/-variánsok zöngességét és zöngességi variációját firtató vélekedések ellentmondásaihoz.

A /h/ zöngésedését illetően a magyar eszközfonetika hőskorában alapvetően két kérdés izgatta a kutatókat. Vajon lehetséges-e egyáltalán a zöngés laringális réshang létrehozása? És ha igen, vajon miként kellene értelmeznünk/elemeznünk azt az igen sajátos zöngét, amely a zöngés laringális réshang képzésére jellemző, és amely réses mozzanatot (glottális rést), valamint zöngés mozzanatot (azaz hangszalagrezgést) egyszerre tartalmaz? Ezt a problematikát is érintő tanulmányában Laziczius (1937) Meyer és Gombocz (1909) munkáját idézi, akik elsőként figyelték meg instrumentálisan, kimoográf segítségével a zöngés [fi]-t a magyarban intervokális helyzetben, sőt más zöngés hangok között is (lásd Laziczius 1937: 306), 20 darab előfordulásban. Laziczius (1937) az idézett szerzők munkája nyomán megállapítja, hogy a magyar szakirodalom már az 1900-as évek elején feltételezi a zöngés laringális réshangot mint a magyar /h/ allofonikus variánsát intervokális (vagy általánosan zöngés hangok közötti) helyzetben (Laziczius 1937: 306–307 állítása szerint ez utóbbi Meyer eredeti megfigyelése és fogalmazásmódja). Emellett pedig arra is rávilágít, hogy a zöngés /h/ létezése két további alapvető kérdést is maga után von. Egyfelől felmerül, hogy vajon ez a zöngés [fi] valóban tekinthető-e a /h/ allofónjának, hiszen a képzéséhez létrehozott rés mérete/jellege és ezzel együtt maga a képzési zörej is alapvetően eltér az alapvariáns [h] ejtésére jellemzőtől. Másfelől pedig kérdéses az is, hogy vajon helyes-e a résmozzanattal együtt járó, különleges zöngét egyszerűen csak „zöngének” tekintenünk, mely terminus alapvetően a hangszalagok többé-kevésbé tökéletes záródásának és nyitódásának váltakozásával járó folyamatot jelöli, mely a további obstruensek vagy magánhangzók képzésekor jellemző.

Az első kérdésben Laziczius (1937) amellett teszi le a voksát, hogy az egyértelmű képzésszerű különbségek ellenére a [fi]-t a /h/ allofónjának kell tekintenünk, hiszen a hangok közti szembenállás nem bír nyelvi relevanciával a

magyarban. Nem meglepő módon a mai napig ez a vélekedés uralja a magyar fonetikai és fonológiai szakirodalmat. (Egyébiránt a [h] és a [ɦ] közötti fonemikus kontraszt a világ nyelveiben általában is igen ritka, 317 nyelv közül mindössze kettőben található meg, vö. UPSID – Maddieson 1984).

Az utóbbi kérdésben Laziczius (1937) a *mormolt zöngé* kifejezés alkalmazására utal (és ezt javasolja), mely éppen a tökéletes zár nélkül képzett zöngehangot jelöli, és különbözteti meg a más obstruensekben és a magánhangzókban alapesetben megfigyelhető teljes vagy modális zöngétől. Manapság erre a zöngeminőségre a magyarban a *mormolt zöngé* terminus mellett a *leheletes zöngé* kifejezést használjuk – ugyanezt a jelenséget pedig az angolban a *glottal murmur*, a *breathy voice* (lásd pl. Maddieson 1984; Stevens 1999), valamint a *whispery voice* (lásd Laver 1994) fedik. Az angol terminusok közül az utóbbi kettő összetett fonációs módot jelöl, amelyben a zöngé nélküli levegőkiáramlás két eltérő típusa, rendre a légzés, illetve a suttogás kombinálódik a zöngéképzés artikulációs gesztusával (lásd Laver 1994; Gobl–Ní Chasaide 1995; Trask 1996). Ennek értelmében tehát az eltérő terminusokat használó szerzők valójában a jelölt zöngeminőség artikulációs és akusztikai paramétereinek leírásában is eltéréseket mutatnak. (A kétféle /h/ ejtéséhez vö. még Esling 2005: 28 leírásait és endoszkópos felvételeit.)

A fentieket összegezve elmondható, hogy a magyar fonetikai szakirodalom már a huszadik század elején feltételezi a zöngés laringális réshangot, és egyúttal fel is ismeri e hang zöngéjének különlegességét. Ugyanakkor a fentiekből az is kitűnik, hogy a /h/ e zöngés allofónjának előfordulási kontextusait ebben az időben még nem tárták fel részleteiben, ahogyan az eszközök adta lehetőségek korlátaiból fakadóan a [ɦ] leheletes zöngéjének akusztikai jellemzőit sem.

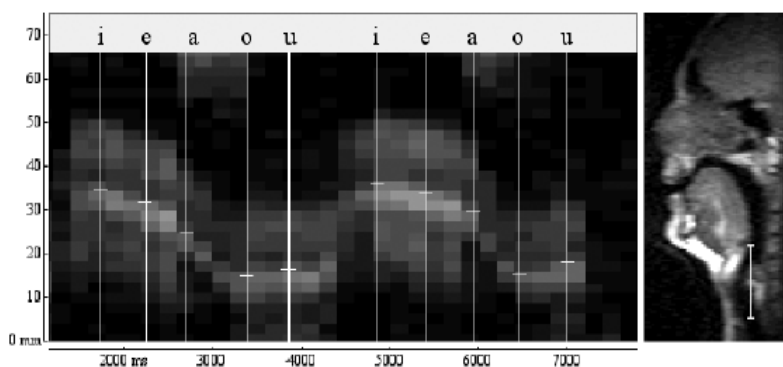
Ahogy arra már utaltunk, a legfrissebb fonetikai kézikönyvekben a /h/ zöngés allofónjával kapcsolatosan vannak ellentmondások. Mégis, ha a feltételezett variánsok képzéshelyétől eltekintünk, és csak az előfordulási környezeteket vesszük figyelembe, viszonylagos egyetértés látszik. A fonológiai szakirodalommal egybehangzó módon ezek a munkák ugyanis a [ɦ]-t előhívó környezetnek tekintik a VhV kontextust (Laziczius 1963/1979; Bolla 1995; Kassai 2005), valamint a szonhV kontextust is (Laziczius 1963/1979; Kassai 1998). Érdekes módon azonban ezek a széles körben elterjedt megállapítások a már említett Meyer és Gombocz (1909) által végzett kimográfós vizsgálaton túlmenően valójában nem támaszkodnak empirikus alapokra (ráadásul sokszor még erre sem hivatkoznak a szerzők), ami azért is problematikus, mert az említett vizsgálat anyagában mindössze húsz, nem szisztematikusan variálódó fonetikai kontextusban megfigyelt /h/-előfordulás volt. Az idézett művön túl pedig mindössze egyetlen olyan kísérlet ismeretes, amely a /h/ zöngésedését vizsgálta több, mint 20 adatponton és a kimográfáznál újabb (és talán némileg megbízhatóbb) eszközökkel. A szonhV kontextusban megvalósuló /h/-k elemzésére ugyanakkor azóta sem született újabb kísérlet, ahogyan

arra a fonológia által javasolt állításra sincs empirikus bizonyíték, mely szerint a hangsúly is szerepet játszana a [h] ~ [h̥] alternációban.

A Meyer és Gombocz által kijelölt utat egyedülként Gósy (2005) munkája követte, aki 50 intervokális helyzetű /h/-előfordulást vizsgált meg a spektrogram és az oszcillogram vizuális elemzésével. Idézett tanulmányában a szerző a /h/ zöngésségét kategorikusan képezte le úgy, hogy a legalább kétharmad részében alacsonyfrekvenciás periodicitást mutató előfordulásokat zöngésként, míg az összes többi előfordulást zöngétlenként elemezte. Mivel a kísérlet exploratív jellegű volt, nyelvi anyagában a /h/-k számos (pontosan nem hivatkozott mennyiségű és minőségű) magyar magánhangzó között álltak (tehát VhV szerkezetekben), és ezek a fonetikai kontextusok változatosak voltak, de nem variálódtak szisztematikusan – előfordultak szimmetrikus és aszimmetrikus, illetve szó belseji és szó határán álló /h/-realizációk is. A /h/ zöngésedésére Gósy (2005) a magánhangzók nyíltságának és előlségének függvényében (is) vont le következtetéseket, bár megjegyzendő, hogy az említett kontextusokban ezek a magánhangzós tulajdonságok (a módszertani fejezetben leírtak – vö. i.m. 9 – és a tanulmánybeli példák alapján feltételezhetően) nem számban kiegyenlített, illetve nem kontrollált összevetésekben szerepeltek (pl. a nyíltság hatásának elemzésénél az összevetett párok nem voltak illesztve az előlség dimenziója mentén). Röviden összefoglalva Gósy (2005) kísérletében arra a következtetésre jut, hogy a /h/ hajlamos zöngésedni akkor, ha a beszédtempó „gyorsabb” (i.m. 11), illetve akkor, ha elől képzett magánhangzók között áll (szimmetrikusan). Bár a kísérletben alkalmazott szólista sajátosságai miatt ezek a konklúziók bizonyos fokú óvatossággal kezelendők, az azonban ennek ellenére is vitathatatlan, hogy az eredmények rámutatnak egy nagyon fontos és kísérletesen alaposabban is ellenőrizendő kérdésre: vajon a /h/-val szomszédos magánhangzók minősége, még pontosabban a magánhangzók két képzési jegye, a nyíltság és az előlség befolyásolják-e a /h/ zöngésedésének tendenciáját? Bár Gósy (2005) beszámolója szerint az előbbire nézve nem talál szignifikáns hatást, eredményei alapján a szomszédos magánhangzók mindkét említett tulajdonsága olyan tényezőnek tekinthető, melyek a /h/ zöngésségére hatással lehet. Az ok, amely miatt úgy látjuk, hogy a nyíltság vizsgálata szintén indokolt, az a tény, amelyet az előlségi hatás magyarázataként is feltételezhetünk: a gége és ezen keresztül a hangszalag-beállítás és a fonáció, valamint a nyelv (mint beszéd szerv) közötti szoros és összetett fiziológiai és aerodinamikai kapcsolat.

Artikulációs vizsgálatok eredményei szerint a (félig) zárt, elől és kerekítés nélkül képzett magánhangzók (pl. /i/, /e/) (de egyes források szerint ide tartozik az elől képzett, nyílt /a/ is, lásd Demolin et al. 2002) ejtésekor a gége magasabb helyzetben van, mint a hátul képzett magánhangzók (pl. /u/, /o/) képzésekor (Hess 1998; Hoole–Kroos 1998; Demolin et al. 2002). (Ez a megállapítás csak azokra a nyelvekre igaz, amelyek a magánhangzók megkülönböztetésében nem élnek az előretolt nyelvgyök [ATR] megkülönböztető jeggyel, vö. Hess 1998.)

Ezt az összefüggést figyelhetjük meg az 1. ábrán, ahol az /i e a o u/ hangsor ejtésekor a gége egyre csökkenő függőleges pozíciója rajzolódik ki az idő függvényében. Noha a magánhangzó-minőség és a gégehelyzet e kapcsolatára utaló eredmények némiképp ellentmondásosak, illetve jelentős, beszélők közötti variációra utalnak (vö. Hoole–Kroos 1998), továbbá nem explicitek az előlég képzési jegy önálló hatására nézvést, mégis alapot szolgáltatnak arra, hogy a következő kiindulásokat fogalmazzuk meg: 1. Az elől képzett magánhangzók képzésekor a gége magasabb helyzetben van, mint a hátul képzettek esetében. 2. Bár az előzőhöz mérten feltehetőleg kisebb mértékben, de a zártabb magánhangzókat szintén elkülöníti a nyíltabbaktól a gége magasabb pozíciója. (Hoole és Kroos 1998 az ajakkerekítés önálló hatását is vizsgálja, de nem sikerül egyértelműen kimutatniuk azt.) Mindez egyúttal azt is jelenti, hogy ezek a magánhangzók a fonáció módjában is eltéréseket mutat(hat)nak, hiszen a függőleges gégehelyzet egyúttal a hangszalagműködésre is hatással van: az alacsonyabb gégehelyzet a hangszalagok abdukcióját (távolítását) vonja magával (Zenker–Zenker 1960; Pabst–Sundberg 1992), míg a magasabb gégehelyzet az addukció (a hangszalagok közelítésének/összeszorításának) mértékét növeli, tehát fokozza a tenziót (Sundberg–Askenfeld 1981; Honda et al. 1999).



1. ábra

A gége függőleges helyzete (mm) az idő (ms) függvényében az /i e a o u/ hangszekvencia kétszeri ejtésekor (férfi beszélő), valós idejű MRI-vel mérve. Az y-tengelyen az alacsonyabb értékek a gége alacsonyabb helyzetére utalnak (Demolin et al. 2002: 554)

Mindezek alapján azt is feltételezhetjük, hogy a magánhangzókra jellemző gégehelyzet és zöngképzés az általuk közrezárt mássalhangzó zöngésségére is kiterjedhet, hiszen az artikulációs gazdaságosság elve szerint a képzésre fordított energiát az emberi szervezet a szükséges minimálisra igyekszik csökkenteni (lásd pl. Lindblom 1990). Így végül arra a feltételezésre jutunk,

hogy a gége valószínűleg a két alacsonyabb gégehelyzetű, azaz hátul képzett és/vagy nyílt magánhangzó között ejtett /h/ képzésekor is alacsonyabb pozícióban van, mint ha a /h/ környezetét elől képzett és/vagy zárt magánhangzók alkotják. Ezt az összefüggést tekinthetjük egyfajta „másodlagos koartikulációnak”, a „másodlagos” kitételrel arra utalva, hogy a koartikuláció itt nem a tradicionális szakirodalom által elsődlegesnek tartott artikulációs gesztusokat vagy jegyeket (például a zöngességet) érinti, hanem olyan másodlagos artikulációs jellemzőket, amilyen például a gége függőleges helyzete (mely csak közvetett módon van hatással például a fonáció módjára is). Következésképpen tehát azt várjuk, hogy a /h/ kontextusát adó magánhangzók minősége, illetve kifejezetten azok nyíltsága és előlsége hatással van a zöngé minőségére és/vagy fenntartására a laringális réshangban.

Kutatásunk átfogó célja a magyar fonetikai és fonológiai irodalom a /h/ zöngésedését érintő állításainak, valamint az azok nyomán felmerült kérdéseknek a tesztelése jól kontrollált, laboratóriumi beszédanyagon. Ennek kezdeti lépéseként a jelen vizsgálatban négy faktor hatását elemezzük a /h/ zöngésedésére: a) az intervokális környezet hatását (szemben a szünet utáni, abszolút szókezdő pozícióval), b) a hangsúly hatását (szemben a hangsúlytalan pozícióval), valamint a szomszédos magánhangzók c) nyíltságának és d) előlségének hatását. A nyelvi anyagban a /h/ előfordulásait kontrolláljuk a szótagbeli pozícióra, és kizárólag szótagkezdetben vizsgáljuk a kérdéses réshangot, mivel fontos, hogy a mássalhangzó képzési helye ne variálódjon az egyes előfordulások között (a szótag onszetjében álló /h/-ra pedig a szakirodalom gyakorlatilag egységesen laringális képzési helyet jósol). Jelen vizsgálatunk fő kérdései a következők:

1. Az intervokális helyzet valóban elicitálja a /h/ zöngésedését szótagkezdetben?
2. Eltérő hatással vannak erre a zöngésedési tendenciára a szomszédos magánhangzók aszerint, hogy azok elől vagy hátul képzettek, illetve nyíltak vagy zártak?
3. Befolyásolja az intervokális /h/ zöngésedését a szótag hangsúlyossága/hangsúlytalansága?

Vélekedésünk szerint az intervokális /h/ zöngésedését elsősorban fonetikai folyamatok határozzák meg, így azt feltételezzük, hogy arra nincs hatással a szótagon megjelenő hangsúly (ahogyan egyébként az esetleges szóhatár sem). Ennek értelmében tehát azt várjuk, hogy az intervokális helyzet minden esetben zöngés realizációkat hív elő, még akkor is, ha az adott *VhV* kontextus szóhatáron átívelően (*V#hV*) jelenik meg, és a /h/-t tartalmazó szótag ezen felül még hangsúlyos is. Ezzel összefüggésben abszolút szókezdő pozícióban a zöngétlen laringális réshang megjelenését jósoljuk. Ugyanakkor a leheletes zöngé sajátos képzéséből és a beszédszervek közötti, a fentiekben vázolt fiziológiai összefüggésekből következően azt is feltételezzük, hogy a magasabb gégehelyzet, tehát az elől képzett és zárt magánhangzós (szimmetrikus) környezet kedvez a [f̥]-szerű realizáció megvalósításának,

azaz a zöngé beindításának és fenntartásának akár a teljes szegmentumon keresztül. A hangszalagokat itt jellemző nagyobb fokú addukció és tenzió ugyanis valószínűsíthetően csökkenti a glottiszban létrehozható rés méretét, amely a zöngétlen, vagy akár a leheletes zöngével képzett hangzók ejtéséhez is szükséges lenne.

E helyütt hangsúlyoznunk kell, hogy vélekedésünk szerint a felvetett kérdések fonetikai szempontok szerint jól kontrollált, felolvasott beszédanyagon, és számban kiegyenlített összevetésekben vizsgálhatók a legmegbízhatóbban, valamint elsősorban olyan paraméterek segítségével, melyek lehetőséget adnak a zöngé graduális természetű megvalósulásának megfigyelésére is. A következőben bemutatásra kerülő kísérletben tehát ezeknek a feltételeknek igyekeztünk megfelelni, és a /h/ zöngésségét a kategorikus (zöngés vs. zöngétlen) elemzési mód helyett a zöngés rész arányával fejezzük ki, illetve a zöngével-zöngésséggel kapcsolatos megfigyeléseinket további, a különbségek gradualitását megragadni képes akusztikai természetű paraméterek segítségével árnyaljuk. Ennélfogva a jelen vizsgálat további célja a magyar /h/-ban megjelenő zöngé minőségének feltáró jellegű, akusztikai elemzése az előbbieken vázolt hatások, de elsősorban a szomszédos magánhangzó minősége, azon belül is kifejezetten a nyíltság és az előlség mentén. Ebből a célból egy olyan akusztikai paramétert elemzünk, amelyet a fonetikában széles körben alkalmaznak (többek között) az obstruensek zöngéjének jellemzésére: a harmonikus-zaj viszony (harmonics-to-noise ratio, HNR) paramétert. A /h/-ra jellemző zöngé minőségét tekintve azt feltételezzük, hogy a szomszédos magánhangzók képzése során (és így feltehetőleg a /h/-ban is) magasabbra és előrébb pozícionált gége a már említett nagyobb fokú hangszalag-addukció és -tenzió révén csökkenti a zöngé leheletességét a [h]-szerű előfordulásokban.

Ahogy az eddigiekben vázoltuk, a feltételezések szerint a magyar laringális réshang szótagkezdetben és intervokálisan megjelenő allofónja zöngés. De ahogyan arra már szintén utaltunk, a [h] zöngéje, a leheletes zöngé merőben eltér más obstruensek, vagy akár a magánhangzók zöngéjétől is. Ahhoz tehát, hogy a [h] intervokális „zöngésségét” vizsgálni tudjuk, olyan paraméterekre van szükségünk, melyek képesek megragadni a leheletes zöngé és a zöngétlenség közti finom (és átmenetekkel járó) akusztikai különbségeket, tehát a zöngé minőségbeli árnyalatait. A leheletes zöngé alább ismertetett artikulációs és akusztikai jellemzői alapján erre a hang harmonicitását számszerűsítő HNR paramétert választottuk ki.

A leheletes zöngé képzése közben a hangszalagok lazán („kis hatásfokkal”) rezegnek, és a záródási szakaszban sosem érik el a teljes glottális zár állapotát. Ennek megfelelően a leheletes zöngé képzése alatt folyamatos a levegőkiáramlás, tehát a zörejképzés is (vö. pl. Laver 1994; Gobl–Ní Chasaide 1995; Stevens 1999). Gobl és Ní Chasaide (1995) szerint az egyik legfontosabb akusztikus kulcs, amely a leheletes zöngét a többi, zörejesebb zöngéképzési módtól, valamint a zöngétlenségtől megkülönbözteti, a zörejösszete-



vőhöz képest dominánsabb periodikus összetevő. Ez alapján úgy látjuk, hogy a /h/ leheletes zöngével és zöngétlenül képzett megvalósulásai közötti különbségeket jól megragadhatja a hang harmonicitását mérő HNR paraméter. Amint azt a neve is sejteti, a HNR (harmonikus-zaj viszony) a periodikus és a zörejes összetevők arányát fejezi ki, tehát a zöngésséget, valamint a zörejséget/zöngétlenséget egyaránt számszerűsíti (Gradoville 2011; G. Kiss 2013) a következő módon. 20 dB HNR-érték azt jelenti, hogy jelben lévő energia 99%-a periodikus, míg 1%-a zaj. Más szóval a 20 dB körüli HNR-értékek szoronás vagy modális/teljes zöngével képzett realizációkra utalnak (ilyet várunk pl. a magánhangzók esetében). Ezzel szemben a 0 dB-es HNR úgy értelmezendő, hogy a hangban lévő energia egyenletesen (fele-fele arányban) oszlik meg a periodikus és zörejes összetevők között (lásd Praat-kézikönyv, Boersma–Weenink 2014b) (a HNR számításának további részleteiért lásd Boersma 1993).

A leheletes zönge másik fontos akusztikai jellemzője az alacsonyabb frekvenciájú összetevők dominanciája a spektrumban a folyamatosan szökő levegő, valamint a zönge mint hangforrás megjelenésének következtében. Ha ugyanis a fonációs ciklus során a hangszalagok nem záródnak tökéletesen, és ezért folyamatos a levegőkiáramlás, a glottiszon átáramló levegő mennyisége sosem esik le nullára, így a levegőáramlás hullámformája megközelítőleg szinuszos lesz. A szinuszos hullámforma pedig nem jelent mást, mint hogy az első harmonikus összetevő, azaz az alapfrekvencia dominálja a létrejött hangot (Stevens 1999). Mindennek tükrében tehát az várható, hogy egy zöngétlen hang spektrális egyensúlya megváltozik, ha abban megjelenik a leheletes (vagy akár a modális) zönge, mégpedig úgy, hogy a spektrum energiaátlaga lefelé tolódik. Éppen ezért feltételezhető lenne, hogy az ún. spektrális súlypont (center of gravity, COG) szintén alkalmas paraméter lehet arra, hogy a zöngétlen és (leheletes) zöngével képzett /h/-variánsok közti, akár átmeneti akusztikai különbségeket, tehát a zönge minőségét leképezze, hiszen ez a mérőszám éppen a spektrumnak az energiával súlyozott átlagfrekvenciáját mutatja, amely a zönge megjelenésével lefelé tolódik, ha a képzési hely és mód jegyek egyébként változatlanok (Gradoville 2011; G. Kiss 2013). Ez az utóbbi azonban nagyon fontos kitétel, mert a COG paraméter érzékeny a frikatíva képzési helyére is, így tehát csak akkor utal megbízhatóan a zönge megjelenésére, amikor a frikatíva képzési helye egyébként nagyjából változatlan. A képzési hely eltolódásával a frikatívában megjelenő zörejjóc és annak spektrális súlypontja (azaz a COG) kísérletes eredmények szerint úgy változik, hogy a szápad elülsőbb részénél képzett (például alveoláris) frikatívák esetén magasabb COG-értékeket mérhetünk, míg az ennél egyre hátrébb képzettek esetében – általánosságban – rendre egyre alacsonyabbat (lásd pl. Gordon et al. 2002 átfogó tanulmányát, melyben hét nyelv zöngétlen frikatíváit hasonlították össze többek közt a COG paraméter és a képzési hely mint elemzési szempont mentén). Mindennek értelmében tehát csak úgy lenne biztosítható, hogy a COG kizárólag a vizsgált frikatíva zöngésségének és/vagy zöngemi-

nőségének akusztikai korrelátumaként legyen értelmezhető, hogy a frikatívát a képzési helyre tökéletesen kontrolláljuk. Ez azonban a /h/ esetében valószínűleg nem lehetséges. Ennek a kérdésnek a kapcsán az alábbiakban röviden ki kell térnünk egy, a /h/ képzését érintő további jellegzetességre is.

A laringális frikatíva /h/-t szájüregi konfigurációjára nézve a fonetikai szakirodalomban általában alulspecifikáltnak tartják (vö. Keating 1988: 282–283; Beckman 1995: 212), ami azt jelenti, hogy azt feltételezzük, hogy a /h/ ejtéséhez csak a gégeszintű artikulációban szükséges bizonyos konfigurációk (a hangszalagrés) megfelelő beállítása, a szupraglottális artikulátorok pozíciója szabadon változtatható. Keating (1998) például ezzel magyarázza azt, hogy a /h/-ban gyakran megfigyelhetők a szomszédos beszédhangokra utaló artikulációs és/vagy akusztikai tulajdonságok, hiszen a /h/ hangsorbeli ejtése közben a szájüregi képzőszervek vagy még az előző beszédhang képzését mutathatják (azaz előrefelé ható/perszeveratív koartikulációs hatást), vagy már a következő beszédhang ejtésére készülhetnek fel (visszafelé ható/anticipációs koartikulációs hatás), avagy egyszerre mindkettő érvényesül. Mi több, állítása szerint az intervokálisan ejtett /h/-k esetében kifejezetten azt tapasztaljuk, hogy a /h/ alatt a magánhangzók képzési helyére utaló második formáns egyszerűen interpolálódik a két magánhangzós cél között, a /h/ tehát nem képez elkülönült artikulációs/akusztikai célt az ejtésben. Mindez arra mutat, hogy az egyébiránt laringális képzési helye ellenére is jelentős eltérések várhatók a /h/ akusztikai szerkezetében (és például kifejezetten a spektrális energiaceloszlásban) a szomszédos magánhangzók minősége mentén. Ugyanakkor arra nézve egyáltalán nem találunk iránymutatást, hogy vajon kell-e eltérést feltételeznünk az intervokális helyzetű alulspecifikált részmáshangzókat és a szünet utáni, szókezdő helyzetben állókat érintő koartikulációs hatások mértékében is ugyanazon minőségű magánhangzók mellett (és ha igen, milyen fokút, jellegűt), pusztán azért, mert az előbbi esetben mindkét irányú koartikuláció, míg a másodikban csak a visszafelé ható koartikuláció hathat. Más szóval tehát nem tisztázott kérdés, hogy a magánhangzó-minőség szerint illesztett #hV és V#hV/VhV típusú hangsorokban a /h/-t érintő koartikulációs hatások következtében várható-e eltérés a nyelv vízszintes és függőleges helyzetét, így például a COG-t illetően az egyébként laringális képzési helyüként specifikált /h/-ban. A fentiek értelmében tehát annak ellenére, hogy szótagkezdő helyzetben laringális képzési helyű frikatívát jósol a szakirodalom, és a COG mérőszám definitíve is kínálkozik a kérdéses hangzó zöngéségének parametrizálására, még magánhangzós kontextusra kontrollált összevetésekben sem biztosított, hogy a COG kizárólag a zöngesség akusztikai korrelátumaként lenne értelmezhető.

#### Kísérleti személyek, anyag, módszer

A kísérletben 19, 22 és 38 év közötti személytől rögzítettünk hanganyagokat (11 férfi, 8 nő, átlagos életkoruk 30 év). A kísérleti személyek mindegyike egy nyelvű, magyar anyanyelvű, ép hallású, és a beszédprodukciónak szerveit

illetően is egészséges beszélő volt. A kísérleti személyeket a felvétel előtt részletesen tájékoztattuk a kísérlet menetéről, majd írásos beleegyező nyilatkozatot kértünk.

Értelmes magyar szavak szótagkezdő pozíciójában előforduló /h/-megvalósulásokat elemeztünk, mely szavakat a kísérleti személyek mondatokba ágyazva olvasták fel. A /h/ három kísérleti kondícióban fordult elő, melyek a magánhangzós kontextus, a szóban elfoglalt pozíció, valamint a mondathangsúly<sup>3</sup> megléte szerint variálódtak:

1. Szó belseji, hangsúlytalan szótagi pozíció (VhV) (ahol a kérdéses hangkapcsolatot tartalmazó szó fókuszpozícióban volt), például: *teher nyomta a vállát, kihív.*

2. Szókezdő, mondathangsúlyos (preverbális fókusz) pozíció (V#hV), például: *keze hevesen kutatott, valaki híreket hozott.*

3. Szó- és megnyilatkozáskezdő (azaz néma szünet utáni), mondathangsúlyos pozíció (#hV), például: *herendi, hirtelen.*

A célszavak válogatását a szószablya webkorpusz segítségével végeztük (Halácsy et al. 2003).

Ebben a kísérleti elrendezésben a #hV kondíció tekinthető kontrollnak, hiszen itt mind a fonetikai, mind a fonológiai szakirodalom laringális képzési helyű zöngétlen frikatívát jósol. A /h/ környezetét képező magánhangzók minden kondícióban az /u ɒ i ε/ voltak, szimmetrikus elrendezésben. Az előlség hatását tehát az /uhu/ és /ɒhɒ/, illetve /ihi/ és /εhe/ csoportok összevetésével, a nyíltság hatását pedig az /uhu/ és /ihi/, illetve /ɒhɒ/ és /εhe/ csoportok összevetésével teszteltük. Megjegyzendő, hogy ilyen módon az előlség és az ajakkerekítés képzési jegyek egyszerre változtak a kondíciók között. Ez azonban a magyar nyelv magánhangzórendszeréből következően (melyben nincs nyílt elől képzett ajakkerekítéses magánhangzó) elkerülhetetlen volt azért, hogy a nyíltság szerint a lehető legszélsőségesebb helyzeteket tesztelhessük. Mivel pedig a hátrébb pozicionált nyelvhez hasonlóan az ajakkerekítés szintén a gége süllyesztését indukál(hat)ja (ha egyáltalán mutat valamiféle szisztematikus hatást, vö. Hoole–Kroos 1998), úgy látjuk, hogy ez a megoldás a lehetőségekhez mérten a legalkalmasabb arra, hogy legalábbis kiegyenlítse a zavaró hatásokat (még ha ezzel esetleg az előlség hatása valamelyest fokozódik is).

Minden stimulust háromszori ismétlésben vettünk fel csendesített szobában, két csatornán: omnidirekcionális mikrofonnal a beszédjelet, elektroglot-

---

<sup>3</sup> Tekintettel arra, hogy a hangsúlyosság hatására vonatkozó forrás (Siptár 1994/2016) nem egyértelműsíti, hogy a hangsúlyosság mely dimenziójáról van szó (szó- és/vagy mondathangsúly), a jelen kutatásba az elméletileg legerősebb hangsúlyt kiváltó pozíciót vontuk be, hiszen ebben az esetben várható a legerősebb hatás is – ha van egyáltalán.

tográffal pedig a hangszalagok mozgását regisztráltuk. (Ez utóbbit a jelen vizsgálatban nem elemizzük.) Összesen 1446 célszón végeztünk akusztikai elemzést.

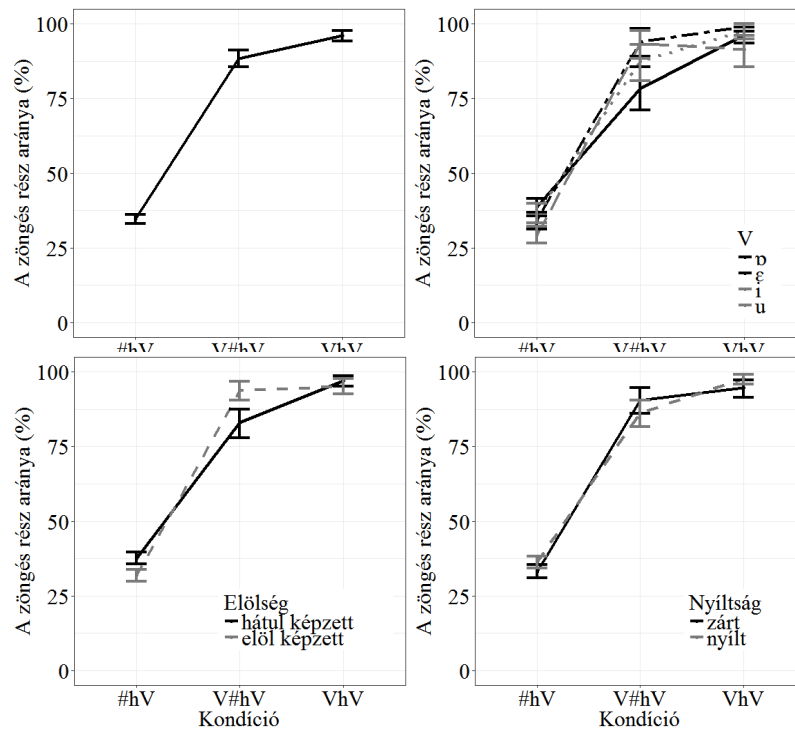
Amint arra már tettünk utalást, vizsgálatunkban a /h/ zöngességét a kategorikus, azaz nominális skálájú változó helyett arányskálára kívántuk leképezni, ezért a /h/ zöngés részének a szegmentum teljes időtartamához viszonyított arányát mértük. Ennek a megbízható megállapításához előkísérletet végeztünk három beszélő felvételein, azaz 450 adatponton. Ebben az előkísérletben elsőként a jelen tanulmány két szerzője címkézte a kérdéses VhV és hV szekvenciákat, majd a /h/ zöngésnek ítélt szakaszát a felvétel 500 Hz-ig szűrt alsó tartományában látható akusztikai kulcsok (elsősorban a jel periodikussága) alapján. A két címkéző által megjelölt zöngés részek arányát összevetve magas korrelációt találtunk ( $\rho = 0,9$ ;  $p < 0,001$ ). Ez után a zöngés rész arányát a Praat (Boersma–Weenink 2014a) voice report funkcióját használva is megállapítottuk, az Eager (2015) által javasolt beállításokkal, azaz az  $f_0$ -tartományt a beszélő nemének függvényében (férfiak: 70–250 Hz, nők: 100–300 Hz) változtattuk. Végül összevetettük a kétféleképpen kinyert adatokat, és az automatikus mérés és az egyik címkéző címkesorából származó adatok között a két címkézőnél található mérhető, erős korrelációt találtunk ( $\rho = 0,7$ ;  $p < 0,001$ ). Ez alapján, valamint a szakirodalomból származó, a miénkkel nagyban egyező eredmények alapján (amelyek a kézi és automatikus elemzés statisztikai ekvivalenciájára utalnak, lásd Eager 2015) a fennmaradó anyagban automatikusan detektáluk a zöngés részt a kézzel szegmentált /h/-realizációkban.

A vizsgálat második céljának értelmében, mely szerint a /h/ zöngéjének minőségi elemzését is el kívántuk végezni a kérdéses kondíciókban, a /h/ előfordulásaiban a már említett HNR paramétert is megmértük, ismét csak a Praat program segítségével. A HNR becsléséhez a teljes /h/ szegmentumot, valamint a Praat alapbeállításait használtuk, kivéve az  $f_0$ -minimumot és -maximumot illetően, amelyeket a nemek szerint variáltunk (ismét csak Eager 2015 irányelveit követve) a nőknél 100 Hz – 300 Hz, a férfiaknál 70 Hz – 250 Hz értéket megadva.

Az adatokat ismételt méréses varianciaanalízissel elemeztük az R programban (R Core Team 2017).

### Eredmények

Elsőként a zöngés rész arányára kapott eredményeket ismertetjük (2. ábra). Ezekben az adatokban általánosságban azt látjuk, hogy az intervokális környezetekben kapott értékek szorosan együtt mozognak, magas zöngearányt mutatva (átlag és szórás a V#hV esetében  $88 \pm 23\%$ , a VhV kontextusban  $96 \pm 13\%$ ), míg a szókezdő helyzetben, kontrollként rögzített /h/-előfordulások zöngességének mértéke ezeknél jóval alacsonyabb (átlag és szórás:  $35 \pm 21\%$ ) (2. ábra, balra fent).



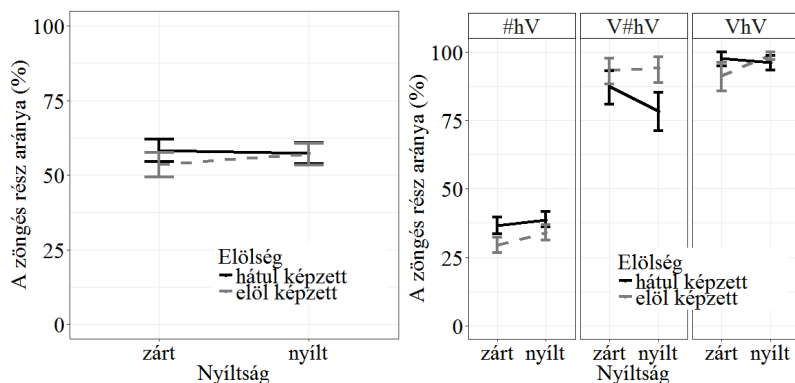
2. ábra

A zöngés rész aránya a /h/-realizációkban: balra fent az összes magánhangzó környezetében együttesen, jobbra fent a magánhangzó-minőség függvényében (/v/: fekete folytonos vonal; /ε/: fekete pontvonal, /i/: szürke szaggatott vonal, /u/: szürke pontozott vonal), balra lent a magánhangzó előlsége és jobbra lent a magánhangzó nyíltsága függvényében (átlag  $\pm$  95% konfidenciaintervallum)

Ha ezeket az adatokat a magánhangzó-minőségek szerinti bontásban elemezzük (2. ábra, jobbra fent), azt látjuk, hogy bár az imént említett tendencia minden magánhangzónál hasonlóan alakul, vannak kisebb eltérések a minőségek mentén, különösen a szókezdő, mondathangsúlyos V#hV pozícióban. Itt ugyanis a hátul képzett /v/ és /u/ mellett álló /h/-ban átlagosan kisebb időarányban jelent meg a zöngé (rendre  $78 \pm 27\%$  és  $87 \pm 24\%$ ), mint az elöl képzett /ε/ és /i/ esetében (mindkét kontextusban  $94 \pm 18\%$ ). Ezen felül kisebb mértékű eltéréseket látunk a szókezdő #hV helyzetben is, ahol az elöl képzett /i/ és /ε/ által közrezárt /h/ mutatkozott zöngétlenebbnek (a zöngés rész ará-

nya itt rendre  $29 \pm 19\%$  és  $34 \pm 20\%$ ) a hátul képzett /u/ és /ɒ/ szomszédságában ejtethez képest (a zöngés rész aránya itt rendre  $37 \pm 22\%$  és  $39 \pm 20\%$ ).

Ha az adatokat a két kérdéses magánhangzó-tulajdonság, a nyíltság és az előlség tükrében elemezzük, azt látjuk, hogy az előlség és a kondíció szignifikáns interakcióban hat azokra ( $F(2, 36) = 8,72$ ;  $p < 0,001$ ; 2. ábra balra lent), ahogyan a nyíltság és az előlség is ( $F(1, 18) = 5,84$ ;  $p < 0,05$ ; 3. ábra; ismételt mérés ANOVA, faktorok: nyíltság, előlség és kondíció). Az első interakció annak a következménye, hogy az elől és hátul képzett magánhangzók csoportja némiképp eltérően viselkedik a V#hV és a #hV kondíciókban: míg a V#hV helyzetben az elől képzettek mutatkoznak valamivel zöngésebbnek (elől képzettek:  $94 \pm 18\%$ , hátul képzettek:  $83 \pm 26\%$ ), ez a viszony a #hV helyzetben megfordul, és bár még az előzőnél is kisebb különbséggel, de az elől képzettek esetében mutat kisebb zöngearányt (hátul képzettek:  $38 \pm 21\%$ , elől képzettek:  $32 \pm 20\%$ ).



3. ábra

A zöngés rész aránya a /h/-realizációkban: bal oldalon a magánhangzó nyíltsága és előlsége, jobb oldalon a magánhangzó nyíltsága és előlsége, illetve a kondíció függvényében (átlag  $\pm$  95% konfidenciaintervallum)

Ez az interakció arra utal, hogy a kondíció (azaz az intervokális kontextus) robusztus hatása némiképp eltérően érvényesül az elől és hátul képzett magánhangzók csoportjában, míg a másik arra, hogy a nyíltság hatása is eltér némiképp e csoportok mentén. Összegzésképp megállapítható tehát, hogy sem a hátul és elől képzett magánhangzók, sem pedig a nyílt és zárt magánhangzók nem látszanak homogén csoportként viselkedni a /h/ zöngés részének arányára gyakorolt hatásukat illetően.

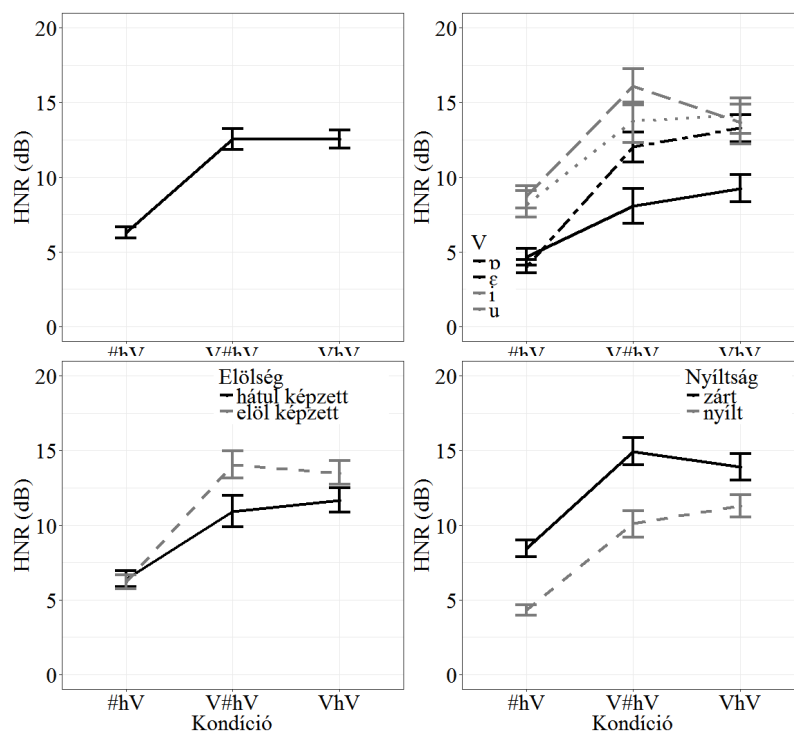
Áttérve a **zöngeminőséget** parametrizáló adatokra lássuk a jel harmonicitását számszerűsítő HNR-értékek alakulását. Ezekben az adatokban általánosságban a zöngés rész arányához hasonló tendenciákat figyelhetünk meg, azaz a kontroll #hV és az intervokális kontextusok egyértelmű szétválását az utób-

biak nagyfokú hasonlósága mellett. A várakozásoknak megfelelően a kontroll kondícióban aperiodikusabb, azaz zörejesebb, míg az intervokális kontextusokban periodikusabb, azaz szonoránsabb /h/-megvalósulásokat látunk (4. ábra, balra fent). A magánhangzó-minőségek szerinti bontás felfedi, hogy ezek a trendek minden magánhangzó esetében megközelítőleg hasonlóan alakulnak (bár a hatás mértéke mutat eltéréseket). Ugyanezen bontás azonban megmutatja azt is, hogy a legzörejesebb /h/-megvalósulásokat az /v/-környezet, míg a legszonoránsabbakat az /i/-környezet váltotta ki (4. ábra, jobbra fent). Egészen pontosan a környezetek következő sorrendje állítható fel a közrezárt /h/ szonoritása szerint: /v/ < /ε/ < /u/ < /i/. Ezen a helyen felhívjuk a figyelmet arra, hogy a HNR tekintetében ismételt az /v/-ra kaptuk a többitől leginkább eltérő adatokat (4. ábra, jobbra fent).

A harmonicitásértékekre az ismételt mérés ANOVA szerint (faktorok: nyíltság, előlség és kondíció) a nyíltság bír szignifikáns hatással ( $F(1, 18) = 51,74$ ;  $p < 0,001$ ), illetve a kondíció az előlséggel interakcióban ( $F(2, 36) = 4,55$ ;  $p < 0,05$ ). Az interakció ebben az esetben minden bizonnyal annak köszönhető, hogy míg a kontroll #hV kondícióban az elől és hátul képzett magánhangzók mellett álló /h/ zöngéssége nem különbözik, addig a két intervokális kontextusban egyértelműen eltérést mutat, mégpedig úgy, hogy az elől képzett magánhangzók kevésbé zörejeseek, mint a hátul képzettek (4. ábra, balra lent). Ugyanakkor azt is megállapíthatjuk, hogy az értékeket befolyásoló fő hatás a /h/ szomszédságában álló magánhangzók nyíltsága olyan módon, hogy a periodikusabb megvalósulások a zártabb magánhangzók környezetében mérhetők. A kapott interakció pedig ugyanúgy, ahogyan a zöngés rész arányánál, ismét csak arra mutat rá, hogy az elől és hátul képzett hangzók, tehát az /v u/ és /ε i/ csoportok valójában heterogén csoportok, legalábbis a /h/ zöngésségére gyakorolt koartikulációs hatásukat illetően.

### Következtetések

A jelen tanulmányban ismertetett akusztikai eredményeink azt mutatják, hogy a laringális képzési helyű, szótagkezdő magyar /h/ zöngésedik intervokális környezetben, és ezt a zöngésedést elsősorban maga az intervokális környezet váltja ki. Ha a /h/ két magánhangzó között áll, megközelítőleg a teljes időtartamának háromnegyedében zöngés lesz, függetlenül attól, hogy az előtte álló magánhangzótól szótaghatár választja-e el úgy, hogy a /h/-t tartalmazó szótag ráadásul hangsúlyos is, valamint szintén függetlenül a /h/ mellett álló magánhangzók minőségétől, illetve azok előlségétől és nyíltságától. Ezzel szemben szünet utáni megnyilatkozás- és szókezdő pozícióban a szakirodalomban foglaltaknak megfelelően a /h/ jóval kisebb, megközelítőleg egyharmad arányban tartalmaz zöngét.



4. ábra

A HNR-értékek a /h/-realizációkban: balra fent az összes magánhangzó környezetében együttesen, jobbra fent a magánhangzó-minőség függvényében (/o/: fekete folytonos vonal; /ε/: fekete pontvonal, /i/: szürke szaggatott vonal, /u/: szürke pontozott vonal), illetve balra lent a magánhangzó elöltsége és jobbra lent a nyíltsága függvényében (átlag  $\pm$  95% konfidenciaintervallum)

Mindebből arra következtethetünk, hogy a feltételezésünk – mely szerint minden intervokális környezet kiváltja a /h/ zöngés realizációját (függetlenül a hangsúly megjelenésétől), ám erre az egyes magánhangzó-minőségek különösen az elöltség és a nyíltság mentén eltérő hatással vannak – részben teljesült, hiszen a magánhangzós képzési jegyek mentén nem láttunk jelentős eltéréseket a zöngés rész arányát tekintve.

Adataink alapján általánosságban az mondható, hogy a szakirodalom alapján zöngésnek jósolt /h/-realizációk megközelítőleg az időtartamuk háromnegyed részében, míg a zöngétlennek jósoltak időtartamuk egyharmadában tartalmaznak a hangszalagrezgésre utaló periodikus összetevőket. Mivel a lehe-



letes zöngével képzett [ɦ] nem kontrasztív a magyarban, ezeknek a határértékeknek a további pontosítása – például percepciós vizsgálatokkal – nem lehetséges. Ennélfogva úgy látjuk, hogy az itt közölt arányok, melyeket nagy mennyiségű, jól kontrollált beszédanyag elemzésével nyertünk, jól jellemzik a tipikus zöngés [ɦ]- és zöngétlen [h]-variánsokat, és ilyenformán referenciaként is szolgálhatnak a jövőben olyan vizsgálatokban, amelyek a /h/ zöngességét kategorikus címkével kívánják leképezni.

A pozíció (szünet utáni vagy intervokális helyzet) és a /h/-val szomszédos magánhangzók elöltségének interakcióban megmutatkozó hatása arra utal, hogy bár valóban minden intervokális /h/ hasonlóan nagy arányban tartalmaz zöngét, a zöngé mennyisége bizonyos mértékben mégis függ a szomszédos magánhangzó minőségétől. Ennek ellenére azonban az intervokális helyzeten belül a magánhangzó-minőségek között látott különbségek a szünet utáni helyzet és az intervokális helyzet között látott különbségekhez képest olyan csekélyek, hogy azok ellenére is nagy biztonsággal zöngésként kategorizálhatnánk minden intervokális, illetve zöngétlenként minden szünet után álló /h/-megvalósulást. Ugyanezen interakcióban megjelenő hatásból pedig egyúttal arra is következtetnünk kell, hogy sem az elől és hátul képzett, sem pedig a nyílt és zárt kategóriák mentén képzett magánhangzó-halmazok nem viselkednek homogén csoportokként a /h/-val kapcsolatban megfigyelhető zöngességi koartikuláció tekintetében. Erre a következtetésünkre a fejezet későbbi pontján, a zöngeminőség számszerűsítésére becsült HNR paraméterre kapott eredmények tárgyalásával kapcsolatosan még visszatérünk.

A zöngeminőség akusztikai mutatójából, a jel harmonicitását megragadó HNR mérőszám értékeiből a zöngé időarányának mutatójánál látottaktól bizonyos tekintetben eltérő kép rajzolódott ki. Ennek kapcsán ugyanis azt láttuk, hogy a legerőteljesebben a szomszédos magánhangzók nyíltsága mentén választotta két, jól elkülöníthető csoportra a /h/ realizációit. A HNR paraméter alapján a pozíció hatása – a zöngés rész esetében látottakhoz hasonlóan – eltérően érvényesült az elől és a hátul képzett csoportok mentén, ez az interakció azonban a zöngé mennyiségénél látottakhoz képest nagyobb eltéréseket okozott. Az eredmények arra mutatnak, hogy a /h/ megvalósulásaiban zárt magánhangzók között – melyek ejtésekor a gége a korábbi eredmények szerint magasabb helyzetben van – a periodikus összetevők dominálják inkább az akusztikai jelet, szemben a nyíltabb (és alacsonyabb gégehelyzettel képzett) magánhangzók mellett álló /h/-változatokkal, amelyek zörejesebbként realizálódnak. Mindezek megerősítik azt a hipotézisünket, mely szerint a gége magasabb helyzete – koartikuláció révén „áttevődve” a résmássalhangzóra – kevésbé zörejes, illetve kevésbé leheletes zöngével megvalósuló /h/-kat hív elő.

A zöngeminőség mutatójánál is látott pozíció  $\times$  elöltség interakció, tehát a pozíciónak az elöltség mentén eltérő hatása ismételtén arra utal, hogy az elől és hátul képzett magánhangzók csoportjaiként bevezetett halmazok az itt vizsgált szempontokból valójában nem homogén csoportok, azaz az /e/ és /o/

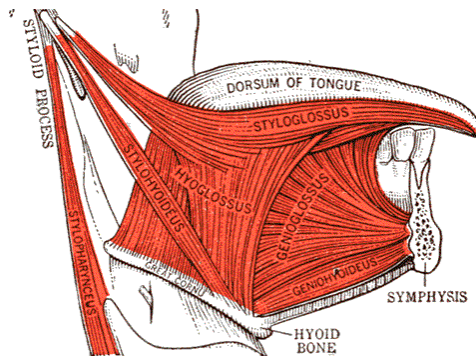
közti különbségek valójában más természetűek, mint az /i/ és /u/ köztié. Mi több, a pozíció és előlség interakcióban lévő hatásából, valamint a szonoritás alapján leginkább úgy tűnik, mintha az itt vizsgált magánhangzók közül mint környezetekből az „elő képzett” és a „hátra képzett” hangok helyett másik két csoport rajzolódna ki, és az /v/ állna szemben az /i u ε/ hangokkal. Ennek az eredménynek az értelmezéséhez érdemes egy rövid kitérőt tennünk a kérdéses magánhangzók hagyományos artikulációs jellemzését illetően (amelyet a jelen tanulmányban mi is alkalmaztunk), és megnéznünk azt, hogy vajon ettől a tradicionális paradigmától eltérően milyen artikulációs gesztusok jellemzik valójában ezeket a beszédhangokat.

A magánhangzókat hagyományosan artikulációs és akusztikai tulajdonságok keverékével, artikulációs, akusztikai és percepciós alapon jellemezzük, a következők szerint. 1. Az állkapocsnyitás szöge/a nyelvállás foka (ti. a nyelv legmagasabb pontjának vagy a nyelvtestnek a függőleges helyzete): e szerint megkülönböztetünk zárt/felső nyelvállású és nyílt/alsó nyelvállású hangzókat, valamint (elsősorban percepciós alapon) két köztes kategóriát, a középzárt/középső nyelvállású és középnílt/alsó nyelvállású hangokét. 2. A nyelv vízszintes helyzete/előlség (ti. a nyelv legmagasabb pontjának vagy a nyelvtestnek a vízszintes helyzete): e szerint beszélünk elő képzett, hátra képzett, illetve mediális/centrális hangokról. 3. Az ajkak működése: e szerint léteznek ajakkerekítéses, valamint ajakréses/kerekítetlen magánhangzók. (Mindezekhez pedig azon nyelvekben, ahol nyelvi jelentősége van, negyedikként a hosszúság fonológiai jegye is társul a jellemzésben.) (Lásd pl. IPA 1999.) Az artikulációt alaposabban megfigyelve azonban egyes vélekedések szerint a magánhangzóképzés e „lingvális-orális modellje” (Esling 2005: 15) meglehetősen leegyszerűsítő.

Elektromiográfias és számítógépes modelleken alapuló vizsgálatsorozatában Honda (1996) a nyelv izmainak koordinációját vizsgálta különös tekintettel a magánhangzók artikulációjára. Eredményei alapján Honda arra a következtetésre jutott, hogy a magánhangzótér négy szélsőséges pozíciójában álló magánhangzójának az ejtéséhez a nyelv négy fő izma párokban rendelődik. Az /i/ ejtésében a genioglossus posterior és genioglossus anterior érintettek, az /æ/ ejtéséhez a genioglossus anterior és a hyoglossus aktivizálódik, az /u/ artikulációjában a genioglossus posterior és a styloglossus játszik szerepet, míg az /a/ létrehozásában a hyoglossus és a styloglossus működnek közre (lásd 5. ábra).

Honda (1996) eredményeit Esling (2005) kiterjeszti a köztes kategóriákra is, és mindebből azt a következtetést vonja le, hogy artikulációs szempontból az általában csak „nyíltsággként” hivatkozott dimenzió valójában alapvetően mást jelent az „elő képzett”, és mást a „hátra képzett” hangok esetében. Míg az „elő képzetteknél” valóban az állkapocs nyitáásszöge tesz különbséget a kérdéses magánhangzó-kategóriák között, addig a „hátra képzetteknél” valójában nyelvemelésről (raising), és nyelvleszorításról (retraction) van szó rendre a „zárt” és „nyílt” „hátra képzett” magánhangzók artikulációjában

(lásd Esling 2005: 23). Mi több, az említett nyelvizmok más-más artikulációs szervekkel való összeköttetése révén Honda (1996) eredményeiből az is következik, hogy a „hátsó képzett” és „nyílt” /v/-szerű hangok fejtik ki a legnagyobb hatást a gége helyzetére, és ebből következően a hangszalagműködésre is a hyoglossus izom, illetve annak a gégével való közvetlen kapcsolata révén (a hyoglossus a nyelvcsontnál ered, így megfeszítése befolyásol(hat)ja a gége függőleges pozícióját).



5. ábra

A nyelv néhány izma jobb oldali oldalnézetből (Clifford Kimber et al. 1893)

Mindezek alapján pedig végül azt is feltételezhetjük, hogy az /v/-szerű hangok fejtik ki a gége függőleges helyzetét érintő legerősebb „másodlagos” (a gége vertikális helyzetét érintő) koartikulációs hatást a velük szomszédos hangok képzésében, tehát feltehetőleg ezek azok a hangzók is, amelyek egyúttal a zöngé minőségét is a leginkább befolyásolják a szomszédos szegmen-tum(ok)ban. A fentiek értelmében úgy véljük, hogy a jelen vizsgálat azon eredménye, amely szerint az /v/, illetve /i/ u /e/ magánhangzócsoporthoz „természetes osztályai” rajzolódnak ki, nem tekinthető váratlannak, mert azt jól megalapozza a kérdéses magánhangzók artikulációs implementációja.

Zárásképpen megfogalmazható, hogy a jelen vizsgálat megerősítette azt a feltételezésünket, mely szerint az intervokálisan ejtett /h/ zöngésedése első-sorban fonetikai folyamat, amelyre kizárólag a magánhangzós környezet bír hatással (koartikuláció révén), a szakirodalomban sugallt fonológiai folyamatok nem. A zöngéesség kérdését finom fonetikai eszközökkel leképező eredményeink ugyanakkor arra is rámutattak, hogy ez a fonetikai folyamat a szomszédos magánhangzók minőségétől, illetve első-sorban a magánhangzók nyíltságától nem függetlenül fejt ki hatását, miközben az elől képzett és hátsó képzett magánhangzók csoportja nem tekinthető homogénnek a nyíltság dimenziója mentén. Ez utóbbi eredményünket a magánhangzók artikulációjának egy, a hagyományos leírásokban használatnál kevésbé leegyszerűsítő ma-

gyarázatával, illetve a nyelv és a gége komplex interakcióinak a figyelembevételével értelmeztük. Terveink szerint a kutatás folytatásaképpen a közeljövőben a fonetikai és fonológiai szakirodalom egy további, a /h/ zöngességét érintő állítását is teszteljük, mely szerint a magánhangzókhoz hasonlóan a /h/ zöngés megvalósulását hívja elő a balról szonoráns, jobbról magánhangzós (szon/hV) környezet is.

### Irodalom

- Beckman, Mary 1995. Implications for phonological theory. In Hardcastle, William – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Cambridge, MA. 199–225.
- Boersma, Paul 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings* 17. 97–110.
- Boersma, Paul – Weenink, David 2014a. *Praat: doing phonetics by computer* [Computer program]. Version 5.4.02. <http://www.praat.org/>. (A letöltés ideje: 2014. december 15.)
- Boersma, Paul – Weenink, David 2014b. *Praat: doing phonetics by computer* [Manual]. Version 5.4.02. <http://www.fon.hum.uva.nl/praat/manual/Harmonicity.html>. (A letöltés ideje: 2014. december 15.)
- Bolla Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Nemzeti Tankönyvkiadó, Budapest.
- Clifford Kimber, Diana – Gray, Carolyn E. – Stackpole, Caroline E. 1893. *Textbook of Anatomy and Physiology*, MacMillan. [http://thediagram.com/3\\_6/some\\_of.html](http://thediagram.com/3_6/some_of.html). (A letöltés ideje: 2017. december 15.)
- Demolin, Didier – Hassid, Sergio – Metens, Thierry – Soquet, Alain 2002. Real-time MRI and articulatory coordination in speech. *Comptes Rendus – Biologies* 325/4. 547–556.
- Eager, C. D. 2015. Automated voicing analysis in Praat: Statistically equivalent to manual segmentation. In The Scottish Consortium for ICPHS 2015 (ed.): *Proceedings of the 18th International Congress of Phonetic Sciences. University of Glasgow, Glasgow, UK*. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0083.pdf>. (A letöltés ideje: 2017. december 15.)
- Esling, John H. 2005. There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics* 50. 13–44.
- G. Kiss, Zoltán 2013. Measuring acoustic correlates of voicing in stops and fricatives. In Szigetvári, Péter (ed.): *VLLXX: Papers Presented to László Varga on his 70th Birthday*. Tinta Könyvkiadó, Budapest. 289–311.
- Gobl, Christer – Ní Chasaide, Ailbhe 1995. Voice source variation. In Hardcastle, William – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Cambridge, MA. 427–461.
- Gordon, Matthew – Barthmaier, Paul – Sands Kathy 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32. 141–174.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.

- Gósy Mária 2005. A /h/ zöngésedése két magánhangzó között. *Beszédkutatás* 2005. 5–20.
- Gradoville, Michael Stephen 2011. Validity in measurements of fricative voicing: Evidence from Argentine Spanish. In Alvord, Scott M. (ed.): *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology. Cascadilla Proceedings Project, Somerville, MA*. 59–74.
- Halácsy Péter – Kornai András – Németh László – Rung András – Szakadát István – Trón Viktor 2003. A szószablya projekt. In Alexin Zoltán – Csentes Dóra (szerk.): *I. Magyar Számítógépes Nyelvészeti Konferencia előadásai (MSZNY 2003)*. 299. [http://eprints.sztaki.hu/7886/1/Kornai\\_1773394\\_ny.pdf](http://eprints.sztaki.hu/7886/1/Kornai_1773394_ny.pdf)
- Hess, Susan A. 1998. *Pharyngeal articulations*. Ph.D. dissertation. UCLA, Los Angeles.
- Honda, Kiyoshi 1996. Organization of tongue articulation for vowels. *Journal of Phonetics* 24. 39–52.
- Honda, Kiyoshi – Hirai, Hiroyuki – Masaki, Shinobu – Shimada, Yasuhiro 1999. Role of vertical larynx movement and cervical lordosis in f0 control. *Language and Speech* 42. 401–411.
- Hoole, Philip – Kroos, Christian 1998. Control of larynx height in vowel production. In Mannell, Robert H. – Robert-Ribes, Jordi (eds.): *Proceedings of the 5th Conference on Spoken Language Processing (ICSLP' 98)* Vol. 2. Sydney. 531–534.
- International Phonetic Association 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press, Cambridge.
- Kassai Iлона 2005. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- Keating, Patricia A. 1988. Underspecification in phonology. *Phonology* 5/2. 275–292.
- Laver, John 1994. *Principles of phonetics*. Cambridge University Press, Cambridge.
- Laziczius Gyula 1937. A zöngés h kérdése. *Magyar Nyelv* 33. 305–310.
- Laziczius Gyula 1963/1979. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht. 403–439.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge University Press, Cambridge.
- Meyer, E[rmst] A[lfred] – Gombocz, Zoltán 1909. *Zur Phonetik der Ungarischen Sprache*. Edvard Berlins Buchdruckerei, Uppsala.
- Pabst, Friedemann – Sundberg, Johan 1992. Tracking multi-channel electroglottograph measurement of larynx height in singers. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 33/2–3. 67–78.
- R Core Team 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Siptár, Péter – Törkenczy, Miklós 2000/2007. *The phonology of Hungarian*. Oxford University Press, New York.
- Siptár Péter 2001. Három felemás magyar mássalhangzó. *Magyar Nyelv* 97. 385–404.
- Siptár Péter 1994/2016. A mássalhangzók. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 183–272. Digitális kiadás: 2016.
- Sundberg, Johan – Askenfelt, Anders 1981. Larynx height and voice source. A relationship? *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 22/2–3. 23–36.

- Stevens, Kenneth N. 1999. *Acoustic phonetics*. MIT Press, Cambridge, MA.
- Szigetvári Péter 1998. Miért nem zöngésedik a [h]? In Büky László – Maleczki Márta (szerk.): *A mai magyar nyelv leírásának újabb módszerei III*. JATE, Szeged. 169–181. [http://seas3.elte.hu/szigetva/papers/miert\\_h.pdf](http://seas3.elte.hu/szigetva/papers/miert_h.pdf) (A letöltés ideje: 2017. december 15.)
- Trask, Robert L. 1996. *A dictionary of phonetics and phonology*. Routledge, London – New York.
- Zenker, Wolfgang – Zenker, Adolf 1960. Über die Regelung der Simmelippenspannung durch von aussen eingreifende Mechanismen. *Folia Phoniatrica* 12/1. 1–36.

### Köszönetnyilvánítás

Köszönjük Varjasi Gergelynek a kísérlet felvételében nyújtott segítségét.

### Intervocalic voicing of the Hungarian /h/

In our research we aim to examine this allophonic alternation of the laryngeal fricative from a phonetic point of view, in an attempt to shed more light on the phonetic and phonological factors that may facilitate or restrain the occurrence of [h̥] in Hungarian, and thus to test previous claims of phonology and phonetics on this issue. As a first step, the present study investigated the effect of two vowel quality features, vowel openness and backness, and a phonological conditioner, pitch-accent on the ratio of voicing that occurs in intervocalic /h/ in laboratory speech. As a secondary aim we also tried to raise questions regarding the very specific type of voice quality this unique fricative exhibits, breathy voice. For this purpose, we also analyzed two more acoustic parameters, center of gravity and the harmonics-to-noise ratio, which are traditionally suggested to reliably and informatively quantify voice quality in fricatives. The results confirmed our hypothesis, that the intervocalic voicing of /h/ may be regarded as a purely phonetic process not governed or restrained by phonological processes, but only by the presence of the two flanking vowels. In a fine-grained phonetic analysis, however, we also showed that this phonetic process may be affected by the vocalic feature, openness of the flanking vowels, via coarticulation, and that front and back vowels may not be regarded as homogenous groups of vowels in the open – close dimension. We interpreted the latter results by taking into account a less simplified explanation of vowel articulation, and the complex interactions of the tongue and the larynx structures.

## A GÉGEMŰKÖDÉS VARIABILITÁSA AZ ÉRZELEMKIFEJEZÉS FÜGGVÉNYÉBEN

**Bartók Márton**

ELTE BTK Fonetikai Tanszék, MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport

### **Bevezetés**

Kutatásomban a magyar *á* mondatszó megvalósulásait vizsgálom. A vizsgálat fő kérdése az, hogy mutat-e eltéréseket az indulatszó fonációjának típusa és alapfrekvenciája annak függvényében, hogy milyen érzelmet fejez ki a beszélő. Elemzem egyrészt az egyes fonációtípusok gyakoriságát a kifejezett érzelem függvényében, másrészt azokban az indulatszókban, ahol fonációtípus-váltást tapasztalok, vizsgálom a különböző fonációtípusok sorrendiségének jellemző mintázatait. Kitérek a nemek közötti eltérésekre is. Az spektrogram és oszcillogram vizuális inspekcója során megállapított fonációtípusokat elektrolottografikus artikulációs adatokkal vetem össze. Vizsgálom az érzelmek hatását a zöngé alapfrekvenciájára is. A mért artikulációs és akusztikai paramétereken főkomponens-elemzést végzek annak megállapítására, vajon az általam alkalmazott érzelemelméleti modell (Russell 1980) dimenziói magyarázzák-e az adatok varianciáját.

A kutatásom fókuszában az emotív kommunikáció áll. Az emocionális, illetve az emotív kommunikáció között először Marty (1908) tett különbséget. Marty szerint az emocionális kommunikáció az érzelmek spontán, önkéntelen felszínre jutása a beszédben. Ezzel szemben az emotív kommunikáció nem automatikus, feltétlen következménye az egyén érzelmi állapotainak, hanem érzelmi, affektív információk célvezérelt kommunikációja a kommunikációs partnerek felé. Az emotív kommunikáció kulturálisan tanult módja annak, hogy érzéseket, attitűdöket fejezzünk ki és értsünk meg különböző interperszonális szituációkban (Stankiewicz 1964). A magyar nyelvészeti szakirodalomban az emotív és az emocionális kommunikáció szembenállása csupán módszertani problémaként, mint a vizsgált érzelmek „természetességét” befolyásoló változó jelenik meg, az automatikus érzelem-felismeréshez szükséges korpuszok építésénél (Sztahó 2013: 31), szemben azzal, hogy az érzelmek beszédbeli megjelenésének ez a két fajtája érzelemelméleti szempontból reflektált módon külön-külön kezelve képezné vizsgálatok tárgyát.

Fontos kérdésfeltevés a szándékos érzelemkifejezés fonetikai kutatásában, hogy milyen nyelvi elemeken vizsgálható az érzelmek hatása kontrolláltan,

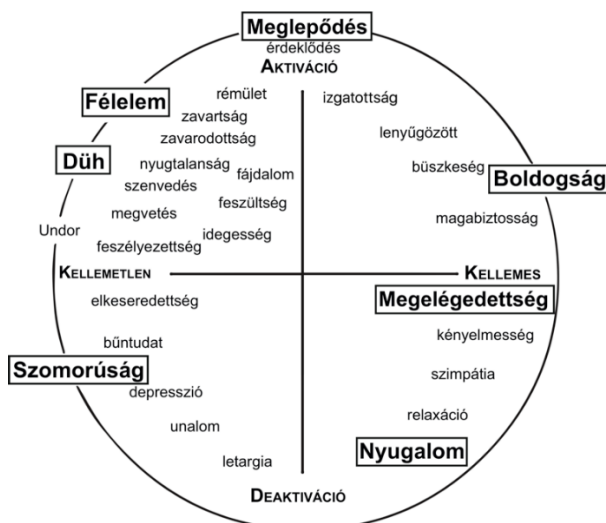
azaz hogyan tartható a nyelvi forma változatlanul, míg a paranyelvi funkció, tehát a kifejezett érzelm változik. A leggyakoribb megoldás az, hogy a kísérlet során az adatközlőknek „érzelmileg semleges értelmű mondatokat” (Sztahó 2013: 23) kell a különböző érzelmeket kifejezve megvalósítaniuk. Ilyen mondatok például a következők: *A falatozóban sört, bort, üdítőitalokat és finom malacsültet lehet kapni.* (Sztahó 2013: 23), *This is not what I expected.* (‘Nem ezt vártam.’ Murray–Arnott 2008: 109), *Heute Abend könnte ich es ihm sagen.* (‘Ma este elmondhatnám neki.’ Burkhardt et al. 2005). Ezek mellett a „semleges értelmű” mondatok mellett a nyelvi anyagot adhatják logatomok (Drioli et al. 2003), illetve tulajdonnevek (Menezes–Maekawa 2006). A fentiek alapján a megszokott módszertani megoldás az érzelmi viszonyulást direkt módon kifejező lexémák, illetve modalitások mellőzése. Ez lehetővé teszi, hogy a paralingvisztikai tartalom szabadon variálható legyen, hiszen nincs olyan nyelvi elem, amely bármely kifejezendő paralingvisztikai tartalomnak ellentmondhatna. Például nem lenne semleges mondatnak tekinthető ezek alapján a *Nagyon örülök a találkozásnak.* mondat, mivel itt a nyelvi tartalom ellentmondana például a FÁJDALOM, SZOMORÚSÁG stb. érzelmek kifejezésének (az ironikus megnyilatkozásoktól eltekintve). A logatomok és a tulajdonnevek esetében szintén ennek az ellentmondásnak az esélye van minimalizálva (bár a tulajdonnevek használata felveti azt a problémát, hogy a kísérleti személyeket esetlegesen saját, az adott tulajdonnévre hallgató ismerőseikhez fűződő érzelmi attitűdje befolyásolhatja). Azonban az, hogy egy adott nyelvi elem nem fejez ki egy adott érzellel ellentétes emocionális tartalmat, tehát hipotetikusán kifejezhet bármilyen más érzelmet, nem jelenti azt, hogy nagy valószínűséggel létezik is olyan kontextus, melyben az adott nyelvi elem az adott érzelmet kifejezve realizálódik.

A jelen kutatásban én éppen ezért egy olyan nyelvi elemet vizsgáltam, mely valóban gyakran fejezi ki érzelmek viszonylag széles palettáját. Ilyen nyelvi elemek például az indulatszók. Ezek a szavak a mondatszók grammatikai kategóriáján belül alkotnak alosztályt. A mondatszók általános jellemzője, hogy „vagy önálló tagolatlan mondatként, vagy szerkesztett tagmondat szintaktikailag szervetlen elemeként fordulnak elő” (Kugler 1998: 341). Az indulatszók olyan mondatszók, melyek a beszélő mentális állapotára vonatkoznak, kifejezhetik az egyén vágyait, kognitív, illetve érzelmi állapotát (Wierzbicka 1992: 164). A beszélő érzelmi állapotát kifejező emotív indulatszók (Wierzbicka 1992: 166) egy része nagyon általános használatú, ezek az indulatszók többféle érzelm kifejezésére alkalmasak. Ilyen funkciógazdag indulatszó például az angol *ah*, amely a FÁJDALOM, MEGLEPŐDÉS, SZÁNALOM, SZENVEDÉLY, PANASZ, NEMTETSZÉS, ÖRÖM és UJJONGÁS kifejezője lehet (Webster 1973), illetve a jiddis *oy*, amelyet Rosten (1968: 274) huszonkilenc különböző érzelm kifejezésére tart alkalmasnak. Funkciógazdag indulatszóként említi Kugler (1998: 341) az *aj*; *jaj*; *hű*; *ó*; *óh*; *ah*; *juj* lexémákat a magyar nyelvben. Dolgozatomban az *á* indulatszó megvalósulásait vizsgál-



tam, mely a fentebbi felsorolásból az *ah* indulatszó mellett, annak alakváltozataként vagy megfelelőjeként jelenik meg a szakirodalomban (vö. Kelemen 1970). Az *á* indulatszót a magyar nyelv értelmező szótára (Bárczi–Ország 1959–1962: 7) szerint többnyire a „rendes [a:] beszédhangnál hosszabban” ejtjük, funkciója pedig a MEGLEPETÉS, CSODÁLKOZÁS, ÁMULAT, TAGADÁS, BOSSZANKODÁS, ELLENKEZÉS, illetve ELHÁRÍTÁS kifejezése.

Szintén dilemmáktól nem mentes kérdés az emotív kommunikációt vizsgáló kutatásokban az, hogy milyen érzelmek kifejezését vizsgáljuk, ugyanis a vizsgálatokban kiválasztott érzelmek kategorizációja és a köztük lévő viszonyrendszer eredendően meghatároz egy érzelmelméleti keretet. Az érzelmek kategorizációjának egy lehetséges módját adja Russel (1980) körmodellje (vö. 1. ábra), mely az aktiváció (függőleges tengely) és a valencia (vízszintes tengely) dimenziója mentén értelmezi a különböző érzelmeket. A modellben az aktiváció az adott érzelmi állapotban jellemző agyi aktiváció (arousal) szintjére, míg a valencia az adott érzelmek negatív vagy pozitív orientációjának mértékére utal.



1. ábra

Russel körmodellje (saját fordítás Zagalo et al. 2005 ábrája alapján). A kutatásban az ábrán bekeretezett érzelmeket vizsgáltam

A körmodell origója a semleges valenciájú, közepes aktivációjú érzelmi állapot, az érzelmek pedig egy origó központú kör pontjaiként helyezkednek el. Dolgozatomban hét érzelmek kifejezését vizsgálom az emotív kommunikációban. Az érzelmek kiválasztása során az volt a cél, hogy az érzelmek körülbe-

lül kiegyenlítőten fedjük le a russeli körmodell kétdimenziós terét. A tanulmányban vizsgált érzelmek az 1. ábrán bekeretezve jelennek meg: MEGLEPŐDÉS, SZOMORÚSÁG, FÉLELEM, MEGELÉGEDETTSÉG, DÜH, NYUGALOM és BOLDOGSÁG. A vizsgált érzelmeket a russeli modell alapján kategorizáltam a valencia és az aktiváció szerint: a kutatásban pozitív valenciájúnak tekintettem a MEGELÉGEDETTSÉGET, NYUGALOM és a BOLDOGSÁG érzelmeket, semleges valenciájúnak a MEGLEPŐDÉST (valamint a kísérletben alkalmazott SEMLEGES kondíciót), negatívnak pedig a SZOMORÚSÁG, FÉLELEM, és DÜH érzelmeket. Alacsony aktivációjúnak tekintettem a SZOMORÚSÁGOT, a MEGELÉGEDETTSÉGET és a NYUGALMAT, magas aktivációjúnak pedig a MEGLEPŐDÉST, BOLDOGSÁGOT, FÉLELMET, és a DÜHÖT.

### A zöngképzés

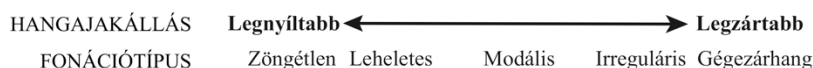
Az érzelmek vokális kifejezéséért Murray és Arnott (1993) szakirodalmi áttekintése alapján a beszédjel három komponense felelős: a fonáció típusát is magába foglaló zöngemínőség, az intonáció és a beszéd temporális vetülete. Dolgozatomban ezek közül a gégeműködés szerepét vizsgálom az érzelm kifejezésben, különös tekintettel a fonációtípusok variabilitására.

A zöngé a hangjak kváziperiodikus rezgése, mely a zöngés hangok esetében a beszédjel forrásaként (vö. Fant 1960) funkcionál. A zöngképzés (fonáció) kezdetekor a hangjak zárt pozícióban állnak, tehát a közöttük lévő rés (glottisz) zárva van. A tüdőből kiáramló levegő ennek következtében feltorlódik a zárt glottisz alatt. Amennyiben a felgyűlt levegő nyomása, az úgynevezett szubglottális nyomás megfelelően nagy, a levegő szétfeszíti a hangjakat, és tovább áramlik a szupraglottális üregrendszerben. Mivel a glottisz szűkületet alkot a gégeben, az itt átáramló levegő sebessége megnő, ez pedig Bernoulli törvényének következtében szívó hatást gyakorol a hangjakra. Ennek következtében a szétfeszített hangjak elkezdnek egymás felé közeledni, majd összezáródnak. A hangjak újbóli zárt állapotából újabb glottális ciklus kezdődhet. Az így létrehozott zöngé különbözhet a zöngéperiódusok gyakorisága, tehát az alaphfrekvencia ( $f_0$ ), valamint az áramló levegő nyomásszintje, tehát az intenzitás nagysága szerint. Ezen a két paraméteren túl jellemezhetjük még a zöngét a fonáció típusával.

Az egyes fonációtípusok Ladefoged (1971) elmélete alapján egy kontinuum mentén helyezkednek el (2. ábra). A kontinuum egyik végpontját a hangjak nyitott, másik végét a hangjak zárt pozíciója jellemzi. Dolgozatomban három fonációtípust különítettem el és elemzek: az irreguláris, a modális, illetve a leheletes zöngét. Az alábbiakban a három fonációtípus képzését a gégeben található izmok feszítettsége és ennek következtében a hangjak relatív helyzete szerinti osztályozás (Laver 1980) alapján ismertetem.

A modális zöngét a hangjak közepes mértékű feszítettsége jellemzi, a glottális periódusok szabályos időközönként követik egymást. Az irreguláris zöngé képzésekor a hangjak mediális kompressziója nagy-, míg longitudi-

nális feszítettségük kismértékű, aminek következtében a tüdőből kiáramló levegő csak ritkábban, és nem feltétlenül szabályos időközönként tudja szétfeszíteni az egymásnak feszülő hangjakakat. A leheletes zöngé képzésekor a hangjakakat egymáshoz közelítő adduktív feszítettség minimális, a longitudinális feszítettség közepes mértékű. Mivel a kismértékű addukció következtében a glottális periódus során a hangjakak soha nem érintkeznek teljes mértékben, a leheletes fonáció során turbulens zöreje keletkezik. A három itt említett fonációtípus közül tehát a hangjakak érintkezése a leheletes zöngé esetében a legkisebb (vö. Kankare et al. 2012; Childers–Lee 1990), az irreguláris fonáció esetén a legnagyobb (vö. Childers–Lee 1990), a modális fonációra pedig közepes mértékű érintkezés jellemző.



2. ábra

A fonációtípusok kontinuumja (saját fordítás Ladefoged 1971 ábrája alapján)

A fonációtípusok funkciói rendkívül sokrétűek a világ különböző nyelveiben. Az irreguláris zöngé kifejezhet fonológiai kontrasztot, jelölheti szegmentumok, szegmentumsorok határát, társalgási egységek lezárását, vagy éppen a beszédjog megtartásának szándékát, illetve lehet szociokulturális szerepe is (bővebben lásd Markó 2013). A leheletes zöngé szintén jelölhet a modális fonációval való fonológiai szembenállást, és lehet szociokulturális szerepe is (Hillenbrand et al. 1994).

A fonációtípusok érzelem-, illetve attitűdkifejező funkcióját több nyelvben is vizsgálták. Laver (1980) impresszionisztikus megfigyelései alapján angol beszélők irreguláris zöngét használhatnak az UNALOM, érdes (nagymértékű általános glottális izomfeszítettséggel jellemezhető) zöngét a DÜH, leheletes zöngét pedig az INTIMITÁS és a BIZALMASSÁG kifejezésére. 34 különböző korábbi akusztikai vizsgálat eredményeit összegezve Scherer (1986) megállapítja, hogy a DÜH, a BOLDOGSÁG és a FÉLELEM esetén jellemzően intenzívebbek a magasfrekvenciás komponensek, amely nagyobb glottális feszítettségre utalhat, a SZOMORÚSÁGRA ezzel szemben alacsonyabb intenzitású komponensek jellemzők a spektrum magasabb frekvenciáin, amely a leheletes zöngé sajátja. Laukkanen és munkatársai (1996) inverz szűréssel vizsgálták logatomok különböző érzelmeket kifejező ejtésében a gégeszintű működés akusztikai vetületét. Eredményeik alapján kisebb a hangjakak érintkezésének aránya a DÜH esetén, és nagyobb arányú érintkezés figyelhető meg a SZOMORÚSÁG, MEGLEPŐDÉS, valamint LELKESEDÉS érzelmeknél. Murray és Arnott (1993) az általuk bemutatott szakirodalom alapján a DÜH és a BOLDOGSÁG érzelmekhez leheletes, a FÉLELEMhez pedig irreguláris fonációt társítanak.

A fentiek alapján elmondható, hogy az érzelmeknek a fonáció típusára kifejtett hatását vizsgáló szakirodalom nem egységes. Ennek hátterében állhat a különböző kutatások eltérő módszertana, az emotív és az emocionális kommunikáció gyakori összemosása, a fonációtípusok leírásának terminológiai sokszínűsége, a vizsgált érzelmek és az érzelmek kifejezéseinek kultúra- és nyelvspecifikus mintázatai. Feltételezhetően a leggyakrabban leírt szembenállás az érzelmi aktiváció szintjével függ össze: számos kutatás eredményei azt mutatják, hogy a magasabb aktivációs szintű érzelmek nagyobb glottális feszítettséggel képzett zöngét eredményeznek, mint az alacsonyabb aktivációjú érzelmek, amelyek esetében inkább leheletesebb zöngé figyelhető meg. A nagyobb agyi aktiváció összefüggése a nagyobb (glottális) izomfeszítettséggel egybeesik az emocionális kommunikáció komponens modelljeiből (pl. Scherer 1986) következő, a spontán érzelmi reakciók fiziológiai jellegzetességeire vonatkozó predikciókkal is. Az aktiváció hatása mellett feltételezhetően vannak különbségek a zöngé minőségében a kifejezett érzelm valenciája függvényében is, azonban Scherer (1986) szerint ezek a zöngéképzés közvetlenebb vizsgálatával, például artikulációs mérésekkel lennének megfigyelhetők. Dolgozatomban ennek megfelelően artikulációs (elektroglottografikus) adatokat is elemzek a fonációtípusok azonosítása és típusokba sorolása mellett. A szakirodalom szerint az agyi aktiváció nem csak a zöngé minőségére, hanem az alaphfrekvenciára is hatással van. Az  $f_0$  érzelmfüggő variabilitását tekintve egységesebb tendencia figyelhető meg a szakirodalomban: a magasabb agyi aktiváció magasabb  $f_0$ -középtértékekkel jár (pl. Davitz 1964, Fónagy 1978, áttekinti Scherer 1986).

A dolgozat célja annak vizsgálata, hogy a magyar *á* indulatszó esetében az [a:] beszédhang mutat-e fonációbeli változatosságot a kifejezett érzelm függvényében. Feltételezem, hogy az érzelmi hatás változatosabb fonációhoz vezet, ezt a feltételezést a mondatzókön belül megjelenő fonációtípusváltások exploratív elemzésével vizsgálom. Feltételezem, hogy összefüggés van a kifejezett érzelm jellemzői (aktiváció és valencia) és az érzelmet kifejező mondatzóban realizálódó [a:] beszédhang fonációjának típusa között, illetve ezzel összefüggésben a kifejezett érzelm és az adott érzelmet kifejező mondatzók zöngéperiódusaiban a hangajkak érintkezésének mértéke között. A nemzetközi szakirodalomban leírtak alapján azt feltételezem, hogy a glottisz nagyobb feszítettsége következtében a hangajkak érintkezésének magasabb aránya, és következésképpen a zártabb hangajakállással képzett irreguláris fonáció lesz jellemző a magas aktivációs szintű érzelmek esetén, ezzel szemben alacsonyabb arányú hangajak-érintkezés, és ezzel összefüggésben leheletes zöngé lesz jellemző az alacsonyabb aktivációs szintű érzelmek esetén. A magasabb aktivációjú érzelmeknél emellett magasabb átlagos alaphfrekvenciát feltételezek, mint alacsonyabb aktiváció esetén. Annak vizsgálatán túl, hogy a russeli körmodell dimenziói mentén mutatkoznak-e különbségek az általam mért változók (alaphfrekvencia, a hangajkak érintkezésének ará-

nya), arra a kérdésre is választ kerestem, hogy vajon ez a két dimenzió (aktíváció és valencia) kellőképpen magyarázza-e a mért függő változók varianciáját, vagy a vizsgált érzelmek más dimenziók mentén, illetve másmilyen csoportokba rendeződnek.

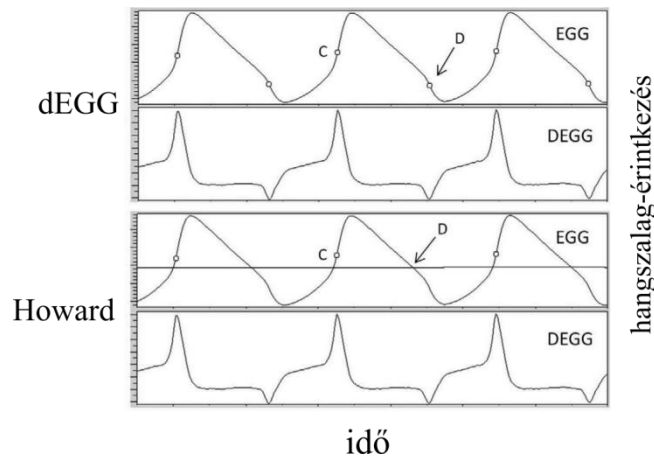
### Anyag, módszer, kísérleti személyek

Bár a kutatás célja az *á* indulatszó érzelmkifejező megvalósulásaink vizsgálata volt, a kísérletben résztvevő adatközlőket az [a:] beszédhang különböző érzelmeket kifejező ejtésére kértem: az adatközlőknek szóló instrukciókban szándékosan nem használtam az *indulatszó* terminust a produkálendő nyelvi egységre, mivel az feltehetőleg nem lett volna érthető, illetőleg összezavarta volna a kísérlet résztvevőit. Az adatközlőket tehát arra kértem, hogy az általam vizsgált érzelmeket kifejezve ejtsenek egy kitartott [a:] hangot. Az adatközlők előtt enyhén döntött monitoron jelentek meg a kifejezendő érzelmek – MEGLEPŐDÉS, SZOMORÚSÁG, FÉLELEM, MEGELÉGEDETTSÉG, DÜH, NYUGALOM, BOLDOGSÁG sorrendben. Az adatközlők azt az utasítást kapták, hogy próbálják minél jobban elképzelni az adott érzelmet, majd minél élethűbben fejezzék ki azt az [a:] hang használatával. Minden érzelem esetén a mondatzó egymást követő három ismétlésére kértem az adatközlőket érzelmeként. A beszédjel rögzítése mellett a zöngképzési folyamatokról elektroglossográf használatával nyertem artikulációs adatokat.

Az elektroglossográf olyan eszköz, amelynek segítségével mérhető a hangajkak érintkezésének mértéke a zöngképzés során. Az elektroglossográf a pajzsporc két oldalára rögzített elektródák között kisfeszültségű, az egészségre teljesen ártalmatlan, magas frekvenciájú elektromos áramot vezet át a glottiszra. Az eszköz jelfeldolgozó egysége a glottisz magasságában átjutó áram feszültségváltozásait rögzíti. Az elektroglossográf által rögzített jelet elektroglossogramnak (EGG) nevezzük. Mivel a hangajkak jobban vezetik az elektromos áramot, mint a gégegében található levegő, a glottisz aktuális vezetőképesége és ennek következtében az átjutó feszültség mértéke a hangajkak érintkezésének mértékétől függ. A jel a minimumát tehát a hangajkak egymástól való teljes eltávolodása, a maximumát pedig a glottisz az adott beszélő esetében jellemző legzártabb állapota esetén veszi fel (Childers–Krishnamurty 1985).

A glottális periódus leírására az EGG többféle jellemzése ismert. A leggyakrabban használt mérőszámok a hangajkak nyitott, illetve zárt szakaszának időtartamarányát adják meg a glottális periódus teljes időtartamához viszonyítva (Rothenberg 1992). A hangajkak érintkezésének arányát a periódusban *contact quotient*-nek (CQ), míg a glottisz nyitott állapotának arányát *open quotient*-nek (OQ) nevezi a szakirodalom. A kutatás során a hangajkak érintkezésének arányát számszerűsíttem, tehát CQ-értékeket mérek. A CQ és az OQ számításának több módja létezik, aszerint, hogy az EGG-jelben milyen kritériumok alapján becsüljük meg a hangajkak érintkezésének és

szétválásának időpontját. Erre a becslésre két főbb módszer, illetve azok kombinációja ismert (Herbst 2004). A módszerek egy része az EGG első deriváltját (dEGG) használja: mivel a derivált jel az eredeti jelben lezajló változások nagyságának, valamint a változások irányának függvényében változik, a hangjakak érintkezésének időpontja feltételezhetően a derivált jel pozitív, szétválása a jel negatív csúcsával (vö. 3. ábra, felső panel) esik egybe (Henrich et al. 2004). Az így kapott nyitódási és záródási időpontok alapján az egymást követő záródási pontok által határolt szakaszokat tekintjük egy periódusnak. A CQ értékét úgy kapjuk meg, hogy a nyitódási (D) és az érintkezési pont (C) közötti időkülönbséget elosztjuk a periódusidővel. A becslés lehetséges még olyan módon, hogy az adott glottális periódusban mért feszültség minimumát tekintjük 0, maximumát 100%-nak, és az érintkezés, valamint a szétválás időpontját ott jelöljük ki, ahol az EGG-jel eléri az általunk kijelölt százalékküszöböt.



3. ábra

A hangszalagok záródásának (C-vel jelölt pont) és nyitódásának (D-vel jelölt pont) meghatározása a dEGG-módszer (fent) és Howard hibrid módszere (lent) esetén (Awan és Awan 2013 Figure 2 alapján, 436)

A módszertani szakirodalomban leggyakrabban 20, 25, 35 és 50%-os, illetve 3/7-es küszöbértéket említenek (Herbst 2004) ami tehát például azt jelenti, hogy a szétválás pillanatát úgy állapítjuk meg, hogy először megállapítjuk a jel lokális maximumát, majd az azt követő első időpillanatot, ahol a jel a maximum érték 3/7-ét veszi fel, és ez utóbbit tekintjük a szétválás pillanatának. Lehetséges továbbá a két módszer kombinált, hibrid alkalmazása is.

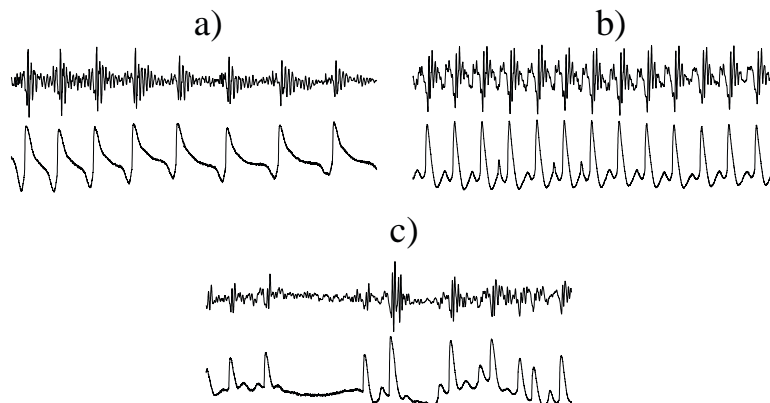
Howard (1995) például a záródás pontját az EGG-jel deriváltjának legnagyobb pozitív csúcsánál, a nyitódás pontját a görbe maximum- és minimum-pontjaihoz viszonyított 3/7-es küszöbénél (Howard 1995) jelöli (vö. 3. ábra, alsó panel). A vizsgálatban a hangajak-érintkezés arányának megállapítására két számítási módot alkalmazok: egyrészt a derivált jel pozitív és negatív csúcsa alapján történő (a továbbiakban  $CQ_{dEGG}$ ), másrészt Howard (1995) hibrid módszerével (a továbbiakban  $CQ_{Howard}$ ) történő becsléseket.

A felvételeket az ELTE Fonetikai Tanszékén végeztem, csendes környezetben. Az akusztikus és elektrolottografikus jelet a Laryngograph Ltd. EGG-D200 típusú elektrolottográf készülékének két csatornáján szimultán rögzítettem. Az elektrolottográf elektródáira konduktív gélt vittem fel. A rögzített hullámformákat (audió és EGG) a Praat programban (Boersma–Weenink 2013) annotáltam: jelöltem egyrészt a mondatszók határait, másrészt a különböző fonációtípusú szakaszokat a mondatszókon belül. A fonációtípusok címkézését az auditív információk mellett az EGG-jel, valamint a beszédjel nyomás-idő függvénye és Fourier-transzformáltja alapján végeztem.

Az EGG-jel elemzését a Peakdet szkript (Michaud–Thi 2007) Praat szkriptnyelven implementált változatában (Kirby 2017) végeztem, mely az eredeti szkripttel ellentétben a dEGG csúcsokat autokorreláció segítségével (Boersma 1993) azonosítja. A mérésekhez az EGG és a dEGG jelet egységesen 10 pontos mozgóátlag-szűrővel simítottam. A periódusok automatikus meghatározásához az alaphérvencia határértékeit nőknél 70–350 Hz-re, férfiaknál 70–250 Hz-re állítottam. Az automatikus elemzés során detektált dEGG csúcsokat manuálisan korrigáltam azokban az esetekben, amennyiben a két becslési módszer esetén az azonos periódusokon mért CQ-értékek jelentősen eltértek, mivel ezek az eltérések általában a csúcsok téves detektálását jelezték. Az így kapott periódusokon mértem a CQ értékét a fentebb ismertetett két módszerrel, az  $f_0$  értékét pedig a periódusidő reciprokaként határoztam meg.

Sok esetben a dEGG-csúcsok helyzetének manuális korrekciója sem volt egyértelműen elvégezhető. Ez a probléma leggyakrabban az irreguláris szakaszok esetén jelentkezett. A mondatszókból az irreguláris fonáció megvalósulásai sokfélék voltak (az irreguláris zöngé néhány, a mondatszókból gyakran megjelenő altípusát a 4. ábra jeleníti meg Keating et al. 2015 kategorizációja alapján). Ez a sokféleség több szempontból is problémát jelentett a CQ-értékek mérésénél. Egyrészt az irreguláris fonáció bizonyos altípusaiban nem állapíthatóak meg zöngperiódusok (aperiodikus irreguláris zöngé, vö. 4.c) ábra), vagy a periódusok két alperiódusra oszthatók (diplofónia, vö. 4.d) ábra). Másrészt, amennyiben mégis megállapítható a CQ értéke, a különböző megvalósulástípusok más-más CQ-értékekhez vezethetnek az annotálás során egyöntetűen irregulárisnak címkézett szakaszokon belül (vö. Keating et al. 2015). A fenti problémákat tekintetbe véve a dolgozat-

ban kizárólag a modális fonációjának annotált periódusokat elemeztem az EGG jel alapján. Így az artikulációs adatok elemzése arra adhat választ, hogy a modálisnak címkézett periódusok mutatnak-e az érzelmek függvényében artikulációs eltéréseket.



4. ábra

Az irreguláris fonáció néhány altípusa egy női beszélő *á* mondatszávaiban: a) prototipikus creak, b) diplofónia, c) aperiodikus irreguláris fonáció. Az ábrák felső regisztrátumán a hangnyomás-idő függvény, az alsó regisztrátumokon az EGG-jel látható

Az adatok statisztikai elemzését az R szoftverrel (R Core Team 2017) végeztem. Kifejezett érzelem és a biológiai nem hatását a fonációtípus-váltást tartalmazó mondatszók összes mondatszóra vetített arányára  $\chi^2$ -próbával elemeztem. A mondatszon belüli fonációtípus-váltások számát a mondatszók fonációbeli komplexitását megragadó numerikus változóként kezeltem, és Pearson-féle korrelációelemzést végeztem a mondatszók így kapott fonációbeli komplexitása és a mondatszók időtartama között. A mondatszon belül elkülönített különböző fonációtípusú szakaszok mondatszóbeli pozíciója és a zöngeszakasz hossza közötti kapcsolatot Spearman-féle rangkorrelációval vizsgáltam. Az adatközlők nemének hatását a nemmodális fonációtípusok gyakoriságára  $\chi^2$ -próbával elemeztem. Az érzelmek, illetve az aktiváció és a valencia hatását a CQ- és az  $f_0$ -értékekre lineáris kevert modellekkel (Kuznetsova et al. 2017) vizsgáltam, az adatközlő személyét random hatásként beállítva. A  $p$ -értékeket Satterthwaite approximációval számoltam, majd páronkénti összevetést végeztem Tukey-féle poszt-hoc teszttel.

A kísérletben mért változókon főkomponens-elemzést (Principal Component Analysis, rövidítve PCA) is végeztem, annak megállapítására,



vajon a russeli modellnek megfelelően a mért adatok varianciáját magyarázza-e az aktiváció és a valencia. A főkomponens-elemzés a változók számának csökkentésére szolgáló statisztikai módszer, a valóban megfigyelt változók közötti kovariancia leírásával. A főkomponens-elemzés eredményeként olyan látens – valójában nem mérhető – változókat (főkomponenseket) tudunk megállapítani, melyek nagymértékben magyarázzák az összes mért változó variabilitását, és százalékosan kifejezhető, hogy az egyes bemeneti változók variabilitása milyen arányban járult hozzá ezeknek a főkomponenseknek a kialakításához. A vizsgálat során a főkomponens-elemzést az általam mért három arányskálájú változón ( $f_0$ ,  $CQ_{dEGG}$  és  $CQ_{Howard}$ ), és az érzelem nominális változón végeztem a numerikus és nominális változókat egyaránt kezelő PCAmixdata (Chavent et al. 2017) R-csomaggal.

A vizsgálatban tízen vettek részt, 5 nő és 5 férfi. A kísérleti személyek életkora 19 és 32 év között szóródik, átlagosan 22,7 év. Bár a szimulált érzelmek vizsgálatában gyakori a professzionális vagy félprofesszionális színészek bemondásainak elemzése, Sztahó (2013: 25) percepció vizsgálatát alapján nem figyelhető meg jelentősebb különbség a szimulált érzelmek felismerésében amatőr beszélők és színészek érzelmi töltetű bemondásainak felismerése között. Ezt figyelembe véve nem törekedtem professzionális színészek beszédének vizsgálatára. Nyolc adatközlő egyetemi hallgató, kettő pedig főállású szolgáltatóipari alkalmazott volt. Mindannyian Budapesten, illetve annak 30 kilométeres körzetében élnek és születtek. Az adatközlőknek a zöngképzést befolyásoló foniatríai természetű problémájuk nem volt. A felvételek elkészítése előtt az adatközlők írásos és szóbeli tájékoztatást kaptak a kísérlet menetéről, az adatok anonim kezeléséről és arról, hogy részvételi bejegyzésüket bármikor, mindenféle hátrányos következmény nélkül visszavonhatják. A tájékoztató után az adatközlők bejegyzési nyilatkozatot töltöttek ki.

### Eredmények

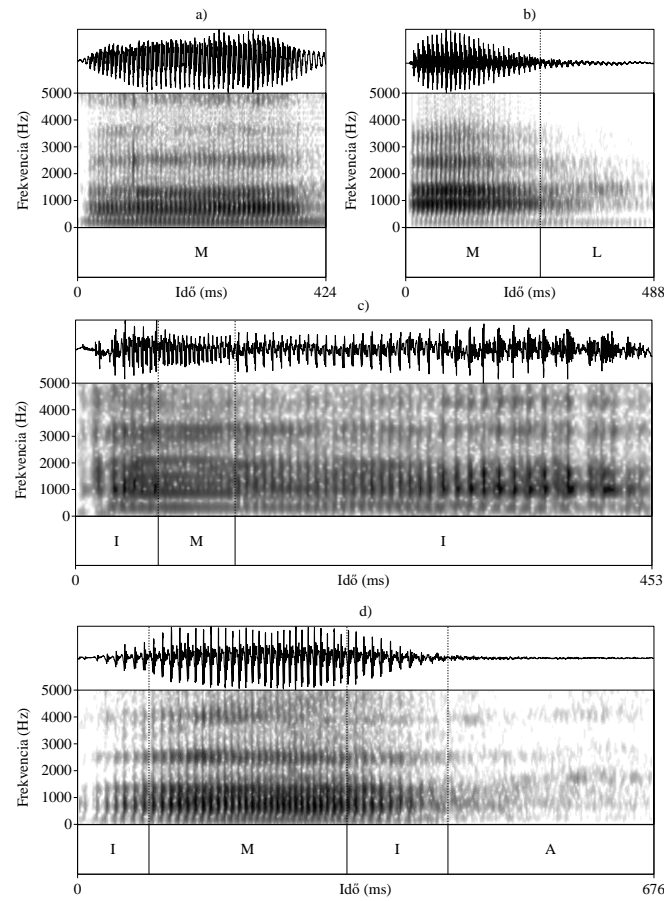
Az adatközlők összesen 244 *a* mondatszót produkáltak, mivel négy adatközlő egy-egy érzelem esetén háromnál több ismétlést valósított meg. A továbbiakban kizárólag a beszélők által ejtett első három ismétlést elemzem minden érzelem esetén. Így összesen 240 mondatszót elemeztem (10 beszélő x 7 érzelem x 3 ismétlés). A mondatszókat átlagos időtartama 481 ms volt, az adatok szórása 155 ms. A legrövidebb mondatszó 135 ms, a leghosszabb 1005 ms hosszúságú volt.

### A manuálisan kategorizált fonációtípusok elemzése

A 240 mondatszón belül összesen 480 szakaszt különítettem el a fonáció típusának elemzése alapján (a dolgozatban a *szakasz*, illetve *zöngeszakasz* kváziterminust ezekre az annotálás során elkülönített egységekre használom). Átlagosan tehát a mondatszók két zöngeszakaszt tartalmaztak (átlagosan egy-

szer változott meg a zöngé minősége egy mondatszó ejtése során). A modális, irreguláris és leheletes fonációjú szakaszok mellett bizonyos esetekben zöngétlen szakaszok is megjelentek a mondatszók részeként, főként a mondatszók végén, annak ellenére, hogy az adatközlőket az [a:] magánhangzó ejtésére kértem. Ezek a zöngétlen szakaszok a zöngétlen laringális spiráns akusztikai szerkezetéhez hasonló, zörejes, átlagosan 125 ms (SD = 79 ms) hosszú intervallumok voltak. A zöngétlen szakaszok egyrészt tekinthetők a mondatszók rövidebbé következtében erőteljesen megjelenő kilégzési zörejeknek. Emellett kapcsolatba hozható a zöngétlen szakaszok megjelenése az indulatszó különböző alakváltozataival. A magyar nyelv értelmező szótára (Bárczi–Ország 1959–1962: 58) szerint az *á* indulatszóval azonos funkcióban jelenik meg az *áh* indulatszó, mely esetében „a rendsz. nagyon hosszan ejtett *á* után többnyire *h* v. *h*-féle hang hallatszik”. Bár a szótárban külön lexémaként jelenik meg az *á* és az *áh*, a két szócikknél felsorolt paralingvisztikai funkciók, tehát a kifejezett érzelmek azonosak, és egyes szerzők alapján nem teljesen egyértelmű, hogy két külön nyelvi egységről van-e szó: Kelemen (1970) kötőjellel, egymás változataiként jelöli az *á-ah* formákat. A kilégzési zörejként való kategorizálás mellett tehát szintén elképzelhető, hogy bizonyos esetekben olyan vokális reprezentációk aktiválódtak a beszélőkben, melyek az indulatszó lexikalizálódott *áh* változatával hozhatók kapcsolatba.

A mondatszók mintegy 69%-a több eltérő fonációtípusú szakaszt is tartalmazott. A mondatszók fonációbeli komplexitása, azaz a fonációtípusváltások száma és a mondatszók időtartama között nem volt megfigyelhető szignifikáns korreláció ( $r = 0,065$ ;  $p = 0,32$ ). Az 5. ábra illusztrálja a mondatszók fonációbeli komplexitásának különböző fokait. A mondatszók 43%-a egy fonációtípus-váltást tartalmazott (vö. 5b) ábra), az adatok 22%-ában kettő (vö. 5c) ábra), 3,3%-ában három (vö. 5d) ábra), 0,8%-ában pedig négy esetben változott meg a fonáció típusa a mondatszón belül.

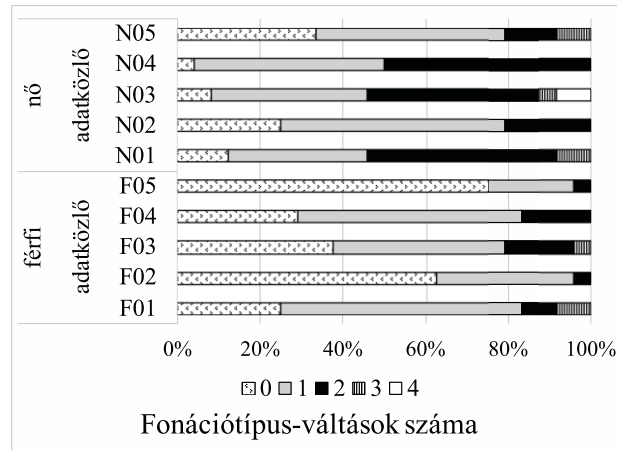


5. ábra

a) Nulla, b) egy, c) kettő, illetve d) három fonációtípus-váltást tartalmazó mondatszók hangnyomásidő-függvénye, hangszínképe és a hozzájuk tartozó címkesor. A fonációtípusok jelölése: irreguláris „I”, modális „M”, leheletes „L”. A zöngétlen szakaszokat „A”-val (aperiodikus) jelöltem

A fonációtípus-váltások száma nagymértékű beszélők közötti variabilitást mutatott. A nulla, egy, két, három, illetve négy fonációtípusú-váltást tartalmazó mondatszók százalékos gyakoriságát adatközlőnként és nemenként szemlélteti a 6. ábra. Volt olyan adatközlő, aki 75%-ban változatlan fonációtípusú mondatszókat produkált, míg olyan is, akinél csupán 4% volt ez az arány. Ez utóbbi egyetlen olyan mondatszót jelent, ahol nem változott a

fonáció típusa. A fonációbeli komplexitás nem csak egyéni jellegzetességet mutat, hanem összefüggésben van a beszélő nemével is. A férfi beszélők a mondatszók 54%-ában változtatták meg legalább egyszer a fonáció típusát, míg ez az arány a nők esetén 83% volt. Ez az eltérés a mondatszók komplexitásának gyakoriságában szignifikáns eltérést mutatott a nemek között [ $\chi^2(1) = 22,419$ ;  $p < 0,001$ ]. Az egy fonációtípus-váltást tartalmazó mondatszók gyakorisága nemenként közel azonos, 42, illetve 43%-os volt. Nagy eltérés figyelhető meg azonban a két váltást tartalmazó mondatszók gyakoriságában: míg a férfiak csak 10%-ban változtatták meg kétszer a fonáció típusát egy mondatszón belül, a nők 34%-ban produkáltak ilyen mondatszókat.

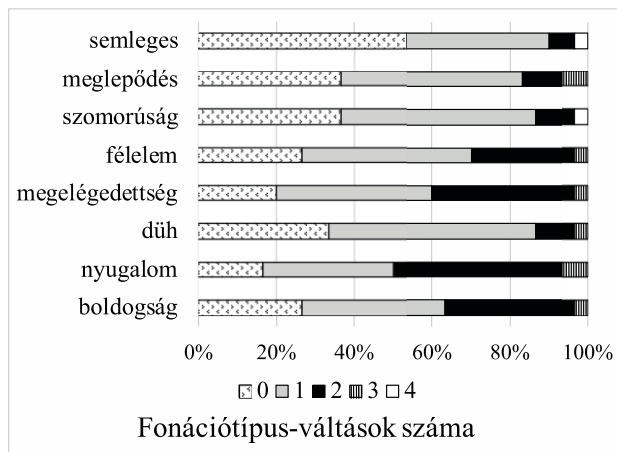


6. ábra

A mondatszókban található fonációtípus-váltások száma (azaz a mondatszók fonációbeli komplexitása) adatközlőnként

A fonációjukat tekintve különböző komplexitású mondatszók százalékos arányát a 7. ábra a vizsgált érzelmek szerint szemlélteti. A SEMLEGES kondícióban ejtett mondatszók 53%-a mutatott variabilitást a fonáció típusában, ezzel szemben az érzelm kifejezés hatására a mondatszók átlagosan 72%-ban legalább egy fonációtípus-váltást tartalmaztak. Ez a különbség a fonációtípusok variabilitásában a SEMLEGES kondíció és az összes érzelm között szignifikáns volt [ $\chi^2(1) = 6,652$ ;  $p < 0,01$ ]. A legnagyobb változatosságot a MEGELÉGEDETTSÉGet, illetve a NYUGALMAT kifejező szakaszok mutatták; ezeknél az érzelmeknél csupán a mondatszók 20, illetve 17%-a mutatott végig egységes fonációt. A mondatszók variabilitásának a SEMLEGES kondícióban mért arányához képest szignifikánsan nagyobb volt a komplex fonációjú szakaszok aránya mind a pozitív valenciájú érzelmek [ $\chi^2(1) =$

8,572;  $p < 0,01$ ], mind az alacsony [ $\chi^2(1) = 7,394$ ;  $p < 0,01$ ] és magas [ $\chi^2(1) = 4,378$ ;  $p < 0,05$ ] aktivációjú érzelmek esetén. A legtöbb érzelm kifejezése esetén az egy fonációtípus-váltást tartalmazó megvalósulások voltak a leggyakoribbak. Kivételt képezett ez alól a NYUGALOM, ami az összes mondatszó 43%-ában két fonációtípus-váltással realizálódott. Az ennél is komplexebb, három, illetve négy váltást tartalmazó mondatszők ugyan kis számban, de az összes kifejezett érzelm esetén megjelentek, még a SEMLEGES kondícióban is.



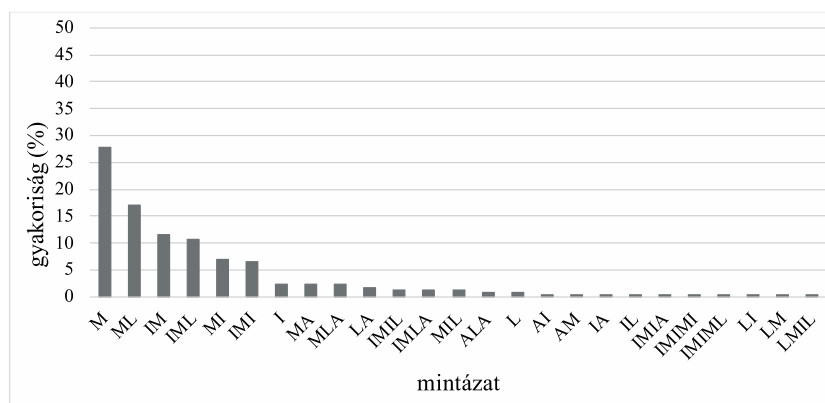
7. ábra

A mondatszőkban található fonációtípus-váltások száma a kifejezett érzelm függvényében

Elemeztem a kategorikus címkékel meghatározott zöngeszakaszok időtartamát. A modális fonációjú szakaszok átlagos hossza 364 ms (a mondatszők átlagos hosszának 76%-a,  $SD = 177$  ms), míg a modálistól eltérő fonációtípusú szakaszok átlagos időtartama ennél jóval rövidebb volt. Az irreguláris szakaszok átlagos időtartama 115 ms (a mondatszők átlagos hosszának 24%-a,  $SD = 139$  ms), a leheletes szakaszoké 158 ms (a mondatszők átlagos hosszának 33%-a,  $SD = 119$  ms), a zöngétlen szakaszoké pedig 125 ms (a mondatszők átlagos hosszának 26%-a,  $SD = 79$  ms) volt.

Elemeztem a mondatszőkat aszerint, hogy a különböző fonációtípusú szakaszok milyen sorrendben követik egymást. A vizsgálati anyagban megjelenő mintázatok gyakoriságát a 8. ábra mutatja be. Az összes adat esetén a leggyakoribb a végig modális megvalósulás volt (28%). A második leggyakoribb eset a modális fonációt követő leheletes zöngképzés (17%), a harmadik és a negyedik leggyakrabban megjelenő mintázat az irreguláris-modális (12%), il-

letve az irreguláris-modális-leheletes (11%) kombináció volt. Emellett az összes mondatszó 7-7 százalékában fordultak elő irreguláris-modális-irreguláris, illetve modális-irreguláris megvalósulások.

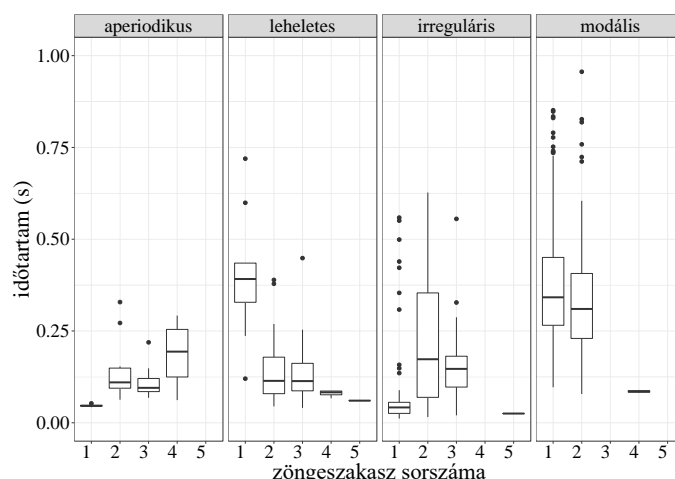


8. ábra

A fonációtípusok sorrendiségének megvalósulásai az összes mondatszóra vetített százalékos gyakorisági aránnyal. A fonációtípusok jelölése: irreguláris „I”, modális „M”, leheletes „L”. A zöngétlen szakaszokat „A”-val (aperiodikus) jelöltem

Vizsgáltam a mondatszón belül betöltött pozíció függvényében a különböző fonációtípusú szakaszok hosszát, ezt a 9. ábra szemlélteti. Mindhárom fonációtípus esetén elmondható, hogy a negyedik, illetve az ötödik pozícióban álló szakaszok átlagosan rövidebbek voltak, mint a mondatszókon belül ennél előbb elhelyezkedő szakaszok. A zöngétlen szakaszokra ezzel szemben éppen a mondatszók végén jellemzőbbek a legnagyobb időtartamértékek. Modális fonációtípusú szakaszok a mondatszókon belül főként első és második pozícióban jelentek meg, ennél a fonációtípusnál gyenge szignifikáns negatív korreláció ( $\rho = -0,16$ ;  $p < 0,05$ ) volt megfigyelhető a zöngeszakaszok pozíciója és időtartamuk között, ami azt jelenti, hogy minél később fordult elő modális zöngeszakasz a mondatszón belül, annál rövidebb volt az időtartama. Az irreguláris szakaszoknál a zöngeszakaszok mondatszóban elfoglalt pozíciója és a szakaszok időtartama közötti korrelációelemzés szignifikáns pozitív korrelációt ( $\rho = 0,51$ ;  $p < 0,001$ ) mutatott, tehát a mondatszóban később megjelenő irreguláris fonációtípusú szakaszok időtartama hosszabb volt. Ezt magyarázhatja az, hogy a mondatszók elején gyakran jelent meg gégezárhang, illetve néhány periódusra kiterjedő irregularitás. A leheletes zöng esetén ezzel szemben a szakaszok időtartama a mondatszó elején volt na-

gyobb, a korrelációelemzés ennél a fonációtípusnál szignifikáns negatív korrelációt ( $\rho = -0,32$ ;  $p < 0,01$ ) mutatott.



9. ábra

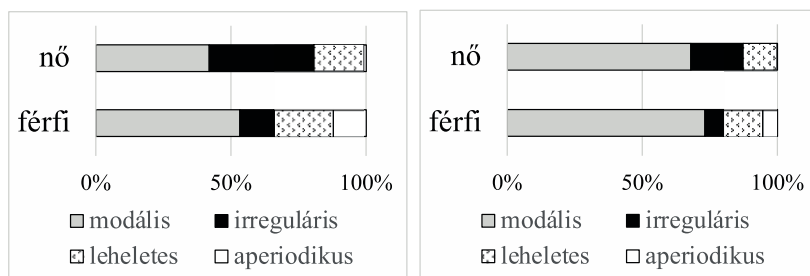
A különböző fonációtípusú szakaszok időtartamértékeinek eloszlása a zöngeszakaszban a mondatszon belüli pozíciójának függvényében

Vizsgáltam a különböző fonációtípusok gyakoriságát, illetve a különböző fonációtípusú szakaszok összhosszához viszonyított arányát is (például az összes modális fonációjú zöngeszakasz időtartamának összege elosztva az összes mondatszó teljes időtartamának szummájával). Mivel a modális fonációjú szakaszok átlagos időtartama mintegy háromszorosa az attól eltérő fonációtípusú szakaszok időtartam-középértékeinek, a fonációtípusok időtartamaránya modális fonáció esetén nagyobb, modálistól eltérő fonáció esetén kisebb értékeket mutat, mint a szakaszok gyakorisági adatai.

A szakaszok 47%-a modális volt, időtartamukat tekintve ezek a szakaszok az összes mondatszó időtartamának 71%-át tették ki. Az ettől eltérő fonációtípusú szakaszok közül az irreguláris fonáció volt a leggyakoribb, ez a szakaszok 27%-át tette ki, ami időtartamát tekintve az összes mondatszó hosszának 13%-át jelentette. A szakaszok 20%-a leheletes zöngével valósult meg, ezen szakaszok együttes hossza 13%-át tette ki az összhosszának. Végül 6%-ban (az összhossz 3%-ában) zöngéképítés nélküli, zörejes szakasz volt észlelhető a mondatszon belül.

Elemeztem a fonációtípusok gyakoriságát (vö. 10. ábra bal oldal) és időtartamarányát (vö. 10. ábra jobb oldal) az adatközlők neme szerint. A nők esetén a szakaszok 58%-a, illetőleg a szakaszok összhosszához 32%-a tért el

a modális regisztertől. A férfiak esetén a szakaszok 47%-a volt a modálístól eltérő fonációtípusú, ezek a szakaszok az összsidőtartam 27%-át jelentették. Az irreguláris zöngével megvalósuló szakaszok gyakorisága a nők esetén 38% (az összsidőtartam 19%-a), a férfiak esetén csupán 13% (az összsidőtartam 7%-a). Az irreguláris zöngével realizálódott szakaszok gyakoriságának a modális szakaszok gyakoriságához viszonyított aránya szignifikáns különbséget mutatott a nemek között [ $\chi^2(1) = 28,037$ ;  $p < 0,001$ ], tehát a nők zöngéje szignifikánsan gyakrabban volt irreguláris. Leheletes zöngét a nők a szakaszok 19%-ában (az összsidőtartam 12%-a), a férfiak pedig a szakaszok 22%-ában (az összsidőtartam 14%-a) produkáltak, ennél a fonációtípusnál nem volt szignifikáns különbség a nemek között a fonációtípus modálishoz viszonyított gyakoriságában. A férfi adatközlők a szakaszok 12%-ában nem képeztek zöngét, ezek a zörejes szakaszok az összsidőtartam 6%-át tették ki. A nők esetében a zöngétlen szakaszok mind gyakoriságuk, mind időtartamarányuk tekintetében 1% alatti értékeket mutattak, a nemek közötti különbség szignifikáns volt [ $\chi^2(1) = 17,310$ ;  $p < 0,001$ ] a modális fonációtípushoz viszonyított gyakoriságra nézve.



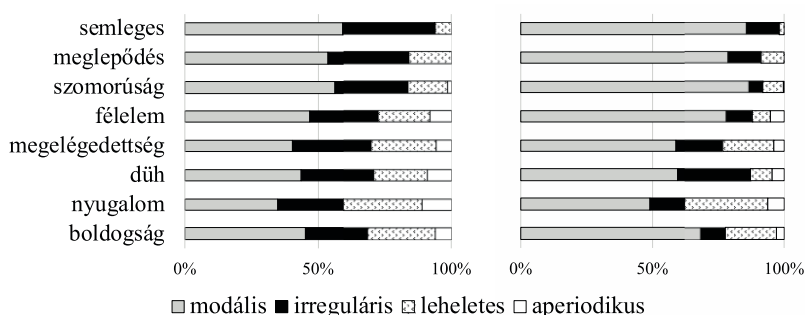
10. ábra

A különböző fonációtípusú szakaszok gyakorisága (balra) és időtartamaránya a mondatszók összsidőtartamához viszonyítva (jobbra) a beszélő nemének függvényében

A mondatszóban megjelenő fonációtípusok és a kifejezni kívánt érzelmek összefüggéseit a fonációtípusok érzelmenkénti gyakorisága és időtartamaránya szerint vizsgáltam. A 11. ábra szemlélteti a fonációtípusok gyakoriságát (balra), illetve időtartamarát (jobbra) a kifejezett érzelem függvényében. Az irreguláris szakaszok gyakorisága nem mutatott nagy eltérést az érzelmek között, a gyakoriságokra kapott értékek 23 és 35% között szóródtak, és a legnagyobb arányban a SEMLEGES kondícióban álló mondatszók tartalmaztak irreguláritást. Az irreguláris szakaszok időtartamának az adott érzelem mondatszóinak összsidőtartamához viszonyított aránya nagyobb variabilitást, és a SEMLEGES mondatszók időtartamarányához (12%) viszonyítva nagyobb elté-



réseket mutat, mint a gyakorisági adatok. A legmagasabb arányban, az összydőtartam 28%-ban a DÜHöt kifejező mondatszók tartalmaztak irreguláris fonációjú szakaszokat. Szintén magas volt az irreguláris fonációjú szakaszok időtartamaránya (18%) a MEGELÉGEDETTSÉG kondícióban álló mondatszók esetén. A NYUGALOM, MEGLEPŐDÉS, FÉLELEM és BOLDOGSÁG kondíciókban álló mondatszók esetében az irreguláris szakaszok időtartamaránya a SEMLEGESES mondatszókra kapott érték körül, 9 és 13% között szóródott. A SZOMORÚSÁG kifejező mondatszók időtartamuk mindössze 5%-ában tartalmaztak irreguláris fonációjú szakaszt. Az érzelmek valenciája szerint nem volt nagy különbség az irregularitás gyakoriságában (negatív valencia esetén 27%, pozitív valencia esetén 26%) és időtartamarányában (negatív valencia esetén 14%, pozitív valencia esetén 13%) Szintén nem befolyásolta az irregularitás gyakoriságát az érzelmek aktivációja, a gyakoriság mind a magas, mind az alacsony aktiváció esetén 27%, az időtartamarány 12, illetve 15% volt.



11. ábra

A különböző fonációtípusú szakaszok gyakorisága (balra) és időtartamaránya (jobbra) a mondatszók összydőtartamához viszonyítva a kifejezett érzelem függvényében

A leheletes fonációjú szakaszok gyakorisága nagy eltéréseket mutatott a mondatszók között. A SEMLEGESES kondícióban álló mondatszók esetén mind a leheletes szakaszok gyakorisága (6%), mind időtartamarányuk (4%) alacsonyabb volt, mint az érzelmeket kifejező mondatszók esetén. A leheletes szakaszok gyakorisága az összes szakasz 15–29%-a között alakult az egyes érzelmek esetén. A leheletes szakaszok időtartamaránya a NYUGALOM kondícióban volt a legmagasabb (32%). Szintén nagy hányadát teszi ki a leheletes fonáció a BOLDOGSÁG és MEGELÉGEDETTSÉG kifejező mondatszók, ezeknél az érzelmeknél a mondatszók időtartamának 19%-át adja. A többi érzelem, tehát a MEGLEPŐDÉS, SZOMORÚSÁG, FÉLELEM és a DÜH esetén a lehe-

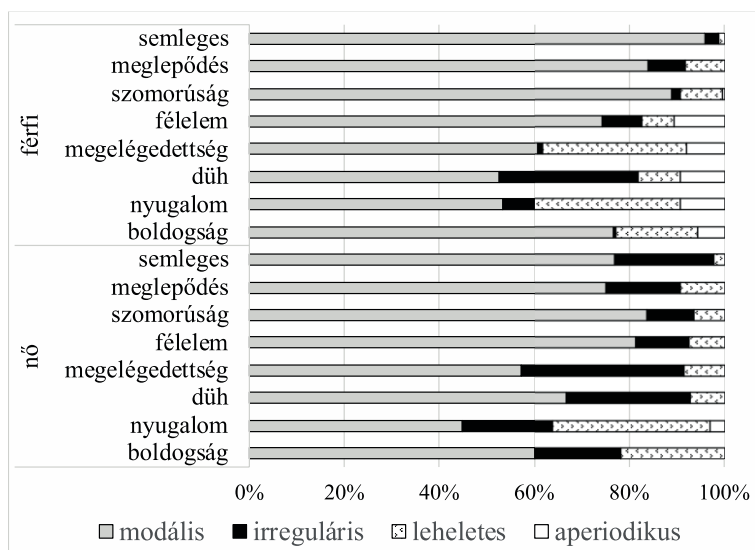
letes szakaszok aránya a SEMLEGES kondícióra jellemző időtartamarány négyesere körül, 7 és 9% között szóródott. A leheletes fonáció gyakrabban, 26%-ban (az összsidőtartam 23%-ában) jelent meg pozitív valenciájú érzelmek esetén, mint negatív érzelmek esetén, itt az irreguláris szakaszok gyakorisága 18%, időtartamaránya 8% volt. Az érzelmek aktivációja szerint kisebb eltérés volt megfigyelhető, az alacsony aktivációjú érzelmek esetén a szakaszok 23%-a (az összsidőtartam 19%-a), pozitív érzelmek esetén a szakaszok 20%-a (az összsidőtartam 11%-a) volt leheletes.

Zöngétlen szakaszok a SZOMORÚSÁG, FÉLELEM, MEGELÉGEDETTSÉG, DÜH, NYUGALOM és a BOLDOGSÁG kondíciókban jelentek meg, 2% és 11% közötti gyakorisággal, 0,3 és 6% közötti időtartamaránnyal. A SEMLEGES kondícióban álló, illetve a MEGLEPŐDÉST kifejező mondatszók egyáltalán nem tartalmaztak zöngétlen szakaszt, a SZOMORÚSÁGOT kifejező mondatszókban pedig egyetlen ilyen intervallumot különítettem el. A zöngétlen és a leheletes szakaszok előfordulási gyakorisága összefüggést mutat: csak azok esetén a mondatszók esetén volt több, mint egy zöngétlen szakasz, amely érzelmeknél a leheletes szakaszok gyakorisága elérte a 19%-ot. Ugyanakkor ezek az érzelmek abban is hasonlóak, hogy a kísérlet során sorrendben később produkálták őket az adatközlők. Lehetséges tehát, hogy a sorrendiség is hatással van a zöngé minőségére, azonban ez nem állapítható meg egyértelműen, hiszen az irreguláris fonáció esetében nem találunk ilyen mintázatot.

Az egyes fonációtípusok különböző érzelmekre jellemző időtartamarányát vizsgáltam az adatközlők neme szerint is (vö. 12. ábra). A nők esetében kiemelkedő volt az irreguláris zöngé időtartamaránya a MEGELÉGEDETTSÉGET (34%), valamint a DÜHÖT (26%) kifejező mondatszók esetében. A MEGLEPŐDÉS, NYUGALOM és a BOLDOGSÁG esetén az irregularitás aránya a SEMLEGES kondícióban álló mondatszókra jellemző időtartamarány (20%) körüli, illetve bizonyos esetekben (SZOMORÚSÁG, FÉLELEM) jóval az alatti volt. A férfiak esetében az irreguláris szakaszok aránya csupán 3% volt a SEMLEGES kondícióban, a SZOMORÚSÁG, a MEGELÉGEDETTSÉG és a BOLDOGSÁG esetén pedig a 2%-ot sem érte el. A NYUGALOM, a MEGLEPŐDÉS és a FÉLELEM esetén az összsidőtartam 6–8%-ában volt irreguláris a zöngé, míg a DÜHÖT kifejező mondatszók esetében ez az arány 29% volt.

A leheletes fonációjú szakaszok aránya a nők esetében 2% volt a SEMLEGES kondícióban, míg a legtöbb nem semleges érzelmek esetében 7% és 9% között alakult a gyakoriság. Ez alól kivételt – ráadásul magasabb értékekkel – a BOLDOGSÁG (22%) és a NYUGALOM (33%) érzelmek képeztek. A férfiak esetén a leheletes zöngé szintén alacsony, 4%-os arányban jelent meg a SEMLEGES kondícióban. A nők adataihoz hasonlóan a legtöbb érzelmek esetén szintén 7 és 9% közötti arányban volt megfigyelhető leheletes fonáció, ennél nagyobb arányban csak a BOLDOGSÁG (17%), a NYUGALOM (31%), és a MEGELÉGEDETTSÉG (30%) esetén jelent meg leheletes zöngé. Zöngétlen szakaszok két megvalósulást leszámítva kizárólag a férfiaknál, a FÉLELEM, MEG-

ELÉGEDETTSÉG, DÜH, NYUGALOM és BOLDOGSÁG kondícióban jelentek meg, az adott érzelmeket kifejező mondatszók összidőtartamának 1–11%-ában.



12. ábra

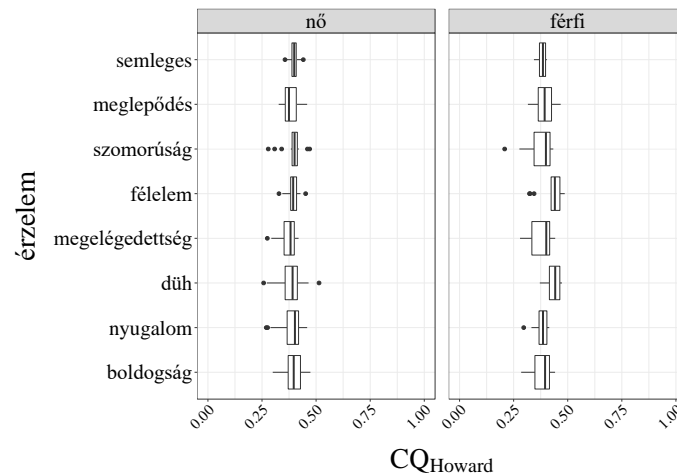
A fonációtípusok időtartamaránya a kifejezett érzelem és az adatközlő nemének függvényében

#### A hangjakak érintkezésének variabilitása a modális zöngével képzett szakaszokban, az érzelem függvényében

A mondatszókban belül modális fonációtípusúnak címkézett szakaszok  $CQ_{\text{Howard}}$ -értékeit a 13. ábra jeleníti meg a kifejezett érzelmek függvényében, nemenként. Nagy különbség volt megfigyelhető mindkét nem esetén a SEMLEGES, illetve érzelmeket kifejező modális szakaszok  $CQ_{\text{Howard}}$ -szórásértékeiben. Míg a SEMLEGES kondícióban a  $CQ_{\text{Howard}}$ -értékek szórása viszonylag alacsony (nőknél 0,03, férfiaknál 0,02) volt, a különböző érzelmek esetében átlagosan kétszer ekkora (nőknél 0,05, férfiaknál 0,04) szórásértékek voltak megfigyelhetők.

Míg a nők  $CQ_{\text{Howard}}$ -értékeit nem befolyásolta a kifejezett érzelem [ $F(7, 103) = 0,673$ ;  $p = 0,695$ ], a férfiak esetében az érzelmek hatása szignifikáns volt [ $F(7, 98) = 5,132$ ;  $p < 0,001$ ]. A férfi adatközlők modális fonációtípusú szakaszainak  $CQ_{\text{Howard}}$ -középértékei átlagosan jóval magasabbak voltak a FÉLELEM (0,43;  $SD = 0,06$ ) és a DÜH (0,44;  $SD = 0,03$ ) esetén, mint a SEMLEGES kondícióban (0,38,  $SD = 0,02$ ), vagy a többi érzelem esetén. A legkisebb  $CQ_{\text{Howard}}$ -értékeket a MEGELÉGEDETTSÉG (0,38;  $SD = 0,06$ ) és a SZOMORÚSÁG

(0,37; SD = 0,07) esetén mértem. A Tukey poszt-hoc teszt alapján a FÉLELEM kondícióban mért értékek szignifikáns eltérést mutattak a SZOMORÚSÁG ( $p < 0,001$ ), a SEMLEGES ( $p < 0,01$ ), NYUGALOM ( $p < 0,01$ ), MEGELÉGEDETTSÉG ( $p < 0,05$ ) és a BOLDOGSÁG ( $p < 0,05$ ) kondíciókhoz képest, emellett szintén szignifikánsan magasabb értékeket mutatott a DÜH a SZOMORÚSÁG ( $p < 0,01$ ), NYUGALOM ( $p < 0,05$ ) és a SEMLEGES ( $p < 0,05$ ) kondíciókkal szembe állítva.



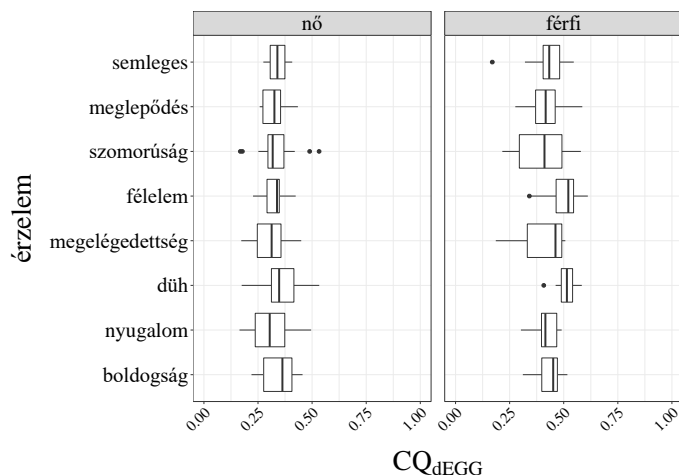
13. ábra

A hangajkak érintkezésének aránya a glottális ciklusban ( $CQ_{\text{Howard}}$ ) a kifejezett érzelm és az adatközlő nemének függvényében a modális fonációtípusúnak címkézett szakaszok esetén (medián és interkvartilis terjedelem)

Vizsgáltam a  $CQ_{\text{Howard}}$  -értékek variabilitását aszerint is, hogy a kifejezett érzelmek a russeli körmodellben milyen valenciával, illetve aktivációval rendelkeznek. A nők  $CQ_{\text{Howard}}$ -értékeire sem az aktiváció [ $F(2, 105) = 0,268$ ;  $p = 0,765$ ], sem a valencia [ $F(2, 105) = 0,845$ ;  $p = 0,433$ ] nem volt hatással. A férfiak esetén a negatív valenciájú érzelmeknél a modális szakaszok  $CQ_{\text{Howard}}$ -átlagértéke 0,41 volt (SD = 0,06), a semlegeseké 0,39 (SD = 0,03), a pozitív valenciájú szakaszoké pedig 0,38 (SD = 0,04). Bár az átlagértékek alapján a pozitív valencia a hangajkak kisebb arányú érintkezésével jár, ez a hatás nem volt szignifikáns sem a három kondíció együttes összevetése [ $F(2, 100) = 2,700$ ;  $p = 0,072$ ], sem a páronkénti összevetés alapján. Az alacsony aktivációjú érzelmek esetén az átlagos  $CQ_{\text{Howard}}$  0,37 volt (SD = 0,05), semleges aktivációnál 0,38 (SD = 0,02), a magas aktivációjú érzelmek esetén pedig 0,41 volt ez az érték (SD = 0,05), tehát az átlagértékek alapján a magasabb aktivációjú érzelmek esetén a hangajkak a glottális periódus nagyobb

hányadában érintkeztek. Ez a hatás szignifikáns volt [ $F(2, 100) = 7,659$ ;  $p < 0,001$ ]. A két független változó között szignifikáns interakciót [ $F(1, 100) = 10,366$ ;  $p < 0,01$ ] találtam: az aktiváció fent leírt hatása csak a negatív valenciájú érzelmek esetén érvényesül, itt az alacsony aktivációjú érzelmeknél 0,37 ( $SD = 0,07$ ), a magas aktivációjúaknál 0,43 ( $SD = 0,05$ ) volt az átlag, míg a pozitív valenciájú érzelmek esetén mind az alacsony, mind a magas aktivációjú érzelmeknél 0,38 volt a  $CQ_{\text{Howard}}$  átlaga ( $SD = 0,05$ ).

A  $CQ_{\text{dEGG}}$  értékeit a modális fonációtípusú szakaszokban a 14. ábra jeleníti meg a kifejezett érzelmek függvényében, nemenként. Az értékek szórása a nők esetén a másik módszerrel mért értékekhez hasonlóan feleakkora volt a SEMLEGES kondícióban (0,04), mint érzelmekifejezés esetén (0,08). A férfiak esetén ilyen eltérés nem volt megfigyelhető, a szórás értéke SEMLEGES kondícióban 0,09, az érzelmeknél átlagosan 0,08 volt.



14. ábra

A hangajkak érintkezésének aránya a glottális ciklusban ( $CQ_{\text{dEGG}}$ ) a kifejezett érzelem és az adatközlő nemének függvényében a modális fonációtípusúnak címkézett szakaszok esetén (medián és interkvartilis terjedeleme)

A nők esetében a  $CQ_{\text{Howard}}$ -értékekhez hasonlóan a  $CQ_{\text{dEGG}}$ -re sem volt szignifikáns hatással a kifejezett érzelem [ $F(7, 103) = 0,989$ ;  $p = 0,443$ ]. A férfi adatközlőknél ezzel szemben szignifikáns hatás volt megfigyelhető [ $F(7, 98) = 4,471$ ;  $p < 0,001$ ]. A legnagyobb átlagos  $CQ_{\text{dEGG}}$ -értékek a DÜH (0,51,  $SD = 0,05$ ) és a FÉLELEM (0,50,  $SD = 0,09$ ) esetén voltak megfigyelhetők, a legkisebb értékek pedig a SZOMORÚSÁGOT (0,40;  $SD = 0,12$ ), a MEGELÉGEDETTSÉGET (0,41;  $SD = 0,11$ ), valamint a MEGLEPŐDÉST (0,41,  $SD = 0,09$ ) kifejező mondatszók esetén. A SEMLEGES kondícióban 0,43 volt az

átlagos érték ( $SD = 0,09$ ). A Tukey poszt-hoc teszttel végzett páronkénti összehasonlításnál pedig ismét a FÉLELEM és a DÜH  $CQ_{DEGG}$ -értékei különböztek szignifikánsan a többi érzelemtől 95%-os konfidenciaszintnél (a FÉLELEM a SZOMORÚSÁGTól, a MEGLEPŐDÉSTől, a NYUGALOMTól és a SEMLEGESTől, a DÜH pedig a SZOMORÚSÁGTól).

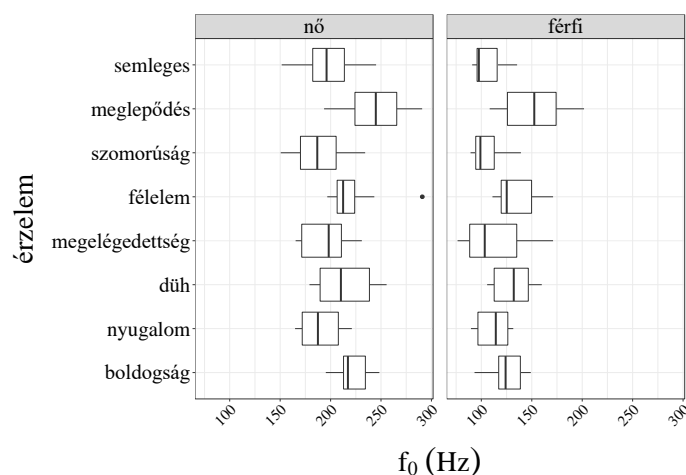
A férfiak  $CQ_{DEGG}$ -értékeire mind a valencia [ $F(2, 100) = 4,750$ ;  $p < 0,05$ ], mind az aktiváció [ $F(2, 100) = 8,129$ ;  $p < 0,001$ ] szignifikáns hatással volt, a két változó között pedig szignifikáns interakció [ $F(1, 100) = 7,361$ ;  $p < 0,01$ ] volt megfigyelhető. Az alacsony, illetve magas aktivációjú érzelmek esetén mért  $CQ$ -átlagok között ismét a negatív valenciájú érzelmek esetén volt megfigyelhető nagyobb különbség, ezeknél az érzelmeknél az alacsonyabb aktivációra mért átlagérték  $0,40$  ( $SD = 0,12$ ), a magasabb aktiváció esetén  $0,50$  ( $SD = 0,07$ ), míg a pozitív valenciájú és alacsony aktivációjú érzelmek  $CQ$ -átlaga ( $0,41$ ;  $SD = 0,08$ ) ennél jóval kevésbé tért el a pozitív valenciájú, magas aktivációjú érzelmek átlagos  $CQ$ -értékétől ( $0,44$ ;  $SD = 0,06$ ).

#### Az alapfrekvencia variabilitása az érzelmek függvényében

A modális periódusokon mért alapfrekvencia értékeit elemeztem az adatközlők neme és a kifejezett érzelmek függvényében. Az  $f_0$ -értékek eloszlását ezek szerint a változók szerint a 15. ábra szemlélteti. A nők átlagos alapfrekvenciája  $209$  Hz ( $SD = 29$  Hz), a férfiaké  $122$  Hz ( $SD = 26$  Hz) volt az összes érzelem esetében. Az érzelmek hatása szignifikáns volt az alapfrekvenciára mind a nők [ $F(7, 103) = 17,768$ ;  $p < 0,001$ ], mind a férfiak [ $F(7, 98) = 17,203$ ;  $p < 0,001$ ] esetén. A legalacsonyabb alapfrekvenciával a nők a SZOMORÚSÁGOT ( $190$  Hz;  $SD = 25$  Hz), és a NYUGALMAT ( $191$  Hz;  $SD = 21$  Hz) fejezték ki, a férfiak esetében a SEMLEGES kondícióban ( $106$  Hz;  $SD = 15$  Hz) és a SZOMORÚSÁG ( $107$  Hz;  $SD = 17$  Hz) esetén mértem a legalacsonyabb alapfrekvencia-értékeket. A legmagasabb  $f_0$  érték mind a férfiak ( $151$  Hz;  $SD = 29$  Hz), mind a nők ( $244$  Hz;  $SD = 28$  Hz) esetén a MEGLEPŐDÉS kifejezésekkor volt mérhető.

Elemeztem a kifejezett érzelem aktivációjának hatását a szakaszok átlagos alapfrekvenciájára. A nők esetén az alacsonyabb aktivációjú érzelmek esetében  $192$  Hz ( $SD = 22$  Hz), a semleges aktivációjúaknál  $196$  Hz ( $SD = 24$  Hz), a magas aktivációjúaknál  $225$  Hz ( $SD = 26$  Hz) volt az átlagos alapfrekvencia, ez az eltérés szignifikáns volt [ $F(2, 106) = 54,551$ ;  $p < 0,001$ ]. A Tukey poszt-hoc teszt mind az alacsony és a magas ( $p < 0,001$ ), mind a semleges és a magas ( $p < 0,001$ ), mind a semleges és az alacsony ( $p < 0,01$ ) aktivációjú érzelmeket kifejező szakaszok  $f_0$ -értékei között szignifikáns eltérést mutatott. A férfi adatközlők esetében az alacsony aktivációjú érzelmeket kifejező szakaszok átlagosan  $110$  Hz ( $SD = 21$  Hz), a semleges aktivációjúak  $106$  Hz ( $SD = 15$  Hz), a magas aktivációjú érzelmek pedig  $136$  Hz-es ( $SD = 24$  Hz) alapfrekvenciával realizálódtak, ez az eltérés szignifikáns volt [ $F(2, 101) = 51,873$ ;  $p < 0,001$ ]. A Tukey poszt-hoc teszt a nők adataihoz hasonlóan ismét

mind az alacsony – magas, mind a semleges – magas, mind a semleges – alacsony szintek között szignifikáns ( $p < 0,001$ ) eltérést mutatott.



15. ábra

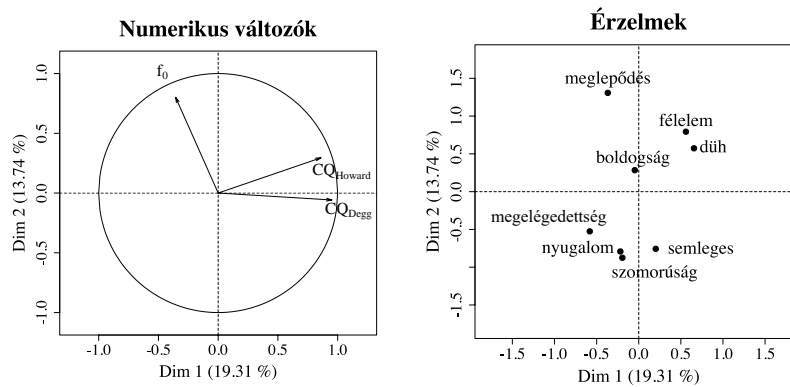
Az alapfrekvencia szakaszonként mért átlagértékei a kifejezett érzelem és az adatközlő nemének függvényében a modális fonációtípusúnak címkézett szakaszok esetén (medián és interkvartilis terjedelem)

Bár összességében a valencia mind a nők [ $F(2, 106) = 13,506$ ;  $p < 0,001$ ], mind a férfiak [ $F(2, 101) = 11,739$ ;  $p < 0,001$ ]  $f_0$ -értékeire szignifikáns hatással volt, a Tukey-féle poszt-hoc teszt alapján mind a nők, mind a férfiak esetében elmondható, hogy a pozitív és negatív valenciájú érzelmek között nem volt megfigyelhető szignifikáns különbség. Az összességében megfigyelt a hatás annak tudható be, hogy a semleges valenciájú érzelmek  $f_0$ -értékei mind a pozitív, mind a negatív érzelmek alapfrekvencia-értékeinél szignifikánsan magasabbak voltak ( $p < 0,001$ ), aminek az az oka, hogy az érzelmek kategorizálásakor a russeli körmodell alapján a MEGLEPŐDÉST semleges valenciájú érzellemnek tekintettem. Mivel pedig a MEGLEPŐDÉSre jellemző kiugróan magas  $f_0$ -értékek nagyban befolyásolták a semleges valenciájú (a MEGLEPŐDÉS mellett csak a SEMLEGES kondíciójú szakaszokat magába foglaló) szakaszok átlagos  $f_0$ -értékeit, valójában nem a valencia, hanem az aktiváció hatása érvényesül.

#### A főkomponens-elemzés eredményei

A főkomponens-elemzés által feltárt első két főkomponenst a 16. ábra jeleníti meg. Az ábrán balra a modell bemeneteként szolgáló arányskálájú változók korrelációjáról kaphatunk képet a két főkomponens által alkotott koor-

dinátarendszerben, míg jobbra ugyanebben a kétdimenziós térben az érzelmek változó nyolc szintjét ábrázoltam. Az első főkomponens (Dim 1) az adathalmaz varianciájának 19%-át magyarázza, és erőteljesen korrelál a  $CQ_{dEGG}$  változóval, valamint kisebb mértékben a  $CQ_{Howard}$  változóval is (ennek a dimenzióknak a kialakításához 47%-ban a  $CQ_{dEGG}$ , 38%-ban a  $CQ_{Howard}$ , 7%-ban az  $f_0$ , és 16 százalékban az érzelmek változó járult hozzá). Ez a főkomponens nagyobb pozitív értékeket mutat a DÜH és a FÉLELEM esetén, kisebb pozitív értéket a SEMLEGES kondíciónál, negatív értéket a BOLDOGSÁG, a SZOMORÚSÁG és a NYUGALOM esetén, kiugróan negatív értéket pedig a MEGLEPŐDÉS és a MEGELÉGEDETTSÉG esetén. Ez az első dimenzió tehát a russeli modell valencia dimenziójának mentén választja szét az egyes érzelmeket, olyan módon, hogy negatív értékeket vesz fel a pozitív valenciájú érzelmek esetén, a negatív valenciájú érzelmek pedig pozitív értékekkel jellemezhetők ebben a dimenzióban. Különbséget az érzelmek megoszlásában a modell és az első dimenzió szerinti pozitív – negatív megoszlás között főként a SZOMORÚSÁG esetében tapasztalunk, ez az érzelmek a pozitív valenciájú érzelmekkel (MEGELÉGEDETTSÉG, NYUGALOM) klasztereződik, ezekhez hasonlóan negatív értéket vesz fel az első dimenzióban. Emellett bár a BOLDOGSÁG esetében az első faktor értéke negatív (-0,047), azonban meglehetősen közel esik a nullához. Viszonylagos eltérés figyelhető meg ezen felül a SEMLEGES kondíciónál is, ami ebben a dimenzióban a modell alapján várt 0 helyett negatív értéket mutat.



16. ábra

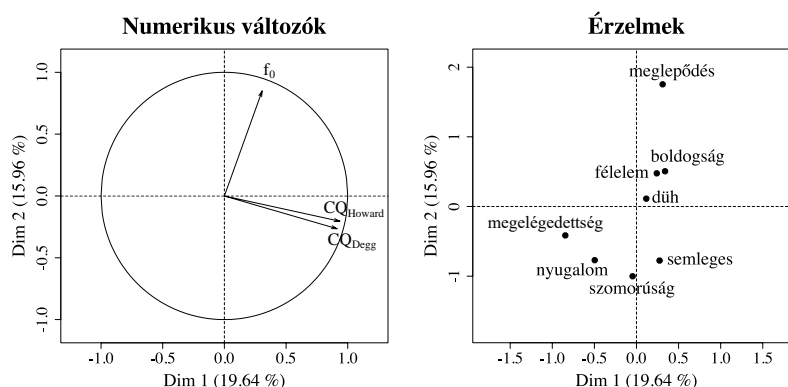
Az adatok varianciáját magyarázó első két dimenzió az összes adatközlőnél

A második főkomponens (Dim 2) a variancia 14%-át magyarázza, és leginkább az alaphétfrekvenciával korrelál (ezt a komponens 47%-ban alakította az  $f_0$ , 0%-ban a  $CQ_{dEGG}$ , 6%-ban a  $CQ_{Howard}$ , és 64%-ban az érzelmek). Ebben a



dimenzióban kiugró értéket mutat a MEGLEPŐDÉS, illetve pozitív értéket tapasztalhatunk még a FÉLELEM, DÜH és a BOLDOGSÁG esetén. Negatív értéket vesz fel ez a látens változó a MEGELÉGEDETTSÉG, NYUGALOM és a SZOMORÚSÁG esetén, valamint a SEMLEGES kondícióban. Ez a főkomponens tehát nagymértékben leképezi a russeli modell aktiváció dimenzióját. Eltérést főként csak a SEMLEGES kondíció esetében tapasztalhatunk: a második dimenzió alapján ezt a kondíciót az alacsony aktivációjú érzelmekhez hasonló módon valószínűzték meg a beszélők.

A nők adatainak főkomponens-elemzésének eredményét ábrázolja a 17. ábra. Az első dimenzió kialakításához az érzelem változó 16%-ban, a második dimenzióhoz 76%-ban járult hozzá, ennek megfelelően az első dimenzió kevésbé rajzolja ki a negatív és a pozitív érzelmek szembenállását, leginkább a MEGELÉGEDETTSÉG és a NYUGALOM negatív értékei különülnek el a többi érzelemtől. A második komponens egyértelműen összefüggésbe hozható az aktivációval: az alacsony aktivációjú érzelmeknél negatív értékeket látunk, a magas aktivációjúaknál az értékek pozitívak (azonban rendkívül kismértékű a nullától való eltérés a DÜH esetében).

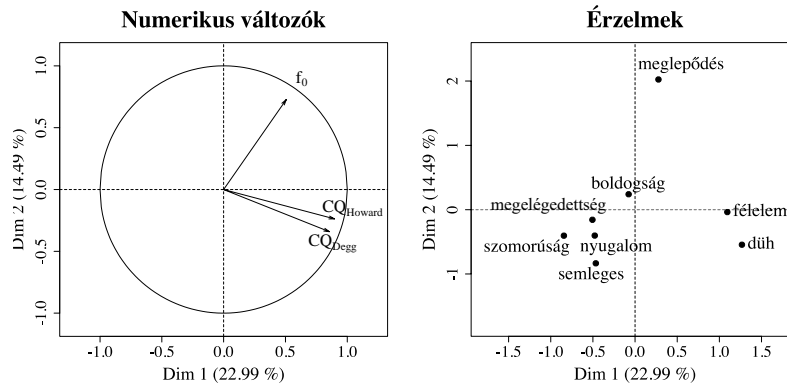


17. ábra

Az adatok varianciáját magyarázó első két dimenzió női adatközlőknél

A férfiak adatainál három jól elkülönülő csoportba klasztereződnek az érzelmek a két komponens által alkotott térben (18. ábra): az adatok variabilitását 23%-ban magyarázó első dimenzió mentén nagy eltérést tapasztalhatunk a kiugró pozitív értékeket mutató FÉLELEM és DÜH és a többi érzelem között (ennek a dimenzióknak a kialakításához 51%-ban járult hozzá az érzelem változó). A többi érzelem a MEGLEPŐDÉS kivételével negatív értékeket vesz fel ebben a dimenzióban, tehát ez a faktor nagyrészt megfeleltethető a valencia dimenziójának – eltekintve a SZOMORÚSÁG kondíciótól, melynél ez a főkom-

ponens a pozitív valenciájú érzelmekre jellemző értéket mutat. A variáciát 14 %-ban magyarázó második dimenzió mentén a kiugróan magas pozitív értéket mutató MEGLEPŐDÉS áll kontrasztban a többi érzellemmel, tehát bár ezt az érzelmet a magas aktiváció jellemzi, a többi magas aktivációjú érzelm nem tér el egységesen ilyen mértékben az alacsony aktivációjúaktól, így ez a faktor kevésbé rajzolja ki az aktiváció dimenzióját (az érzelm változó 75%-ban járult hozzá a második dimenzióhoz).



18. ábra

Az adatok variációját magyarázó első két dimenzió férfi adatközlőknél

### Következtetések

A jelen tanulmányban a zöngképzés és az alaphérfvencia variabilitását elemeztem a magyar á mondatszó realizációiban, emotív kommunikáció során, azaz szándékolt érzelmközlésekben. Kutatásomban az adatközlőket arra kértem, hogy egy kitartott [a:] hang segítségével fejezzenek ki különböző érzelmeket. A kísérlet során produkált egységeket indulatszónak tekintettem, hiszen formájuk és funkciójuk tekintetében is megfelelnek az emotív indulatszó prototipikus példányainak.

Az indulatszók fonációtípusának elemzése során azt tapasztaltam, hogy egy indulatszóban akár négyszer is változhatott a fonáció típusa. A SEMLEGES kondícióban produkált érzelmek esetén ez a variabilitás szignifikánsan kisebb volt, mint a különböző érzelmeket kifejező mondatszók esetében. A mondatszók fonációja tehát változatosabb lesz az érzelmkifejezés hatására.

Az irreguláris szakaszoknak az összes mondatszó időtartamához viszonyított időtartamaránya nem mutatott eltéréseket a kifejezett érzelm valenciája (pozitív-negatív orientációja) szerint, azonban volt néhány érzelm, amely esetén kiugró volt az irreguláris szakaszok időtartamaránya: a mondatszók összéidőtartamának férfiaknál 29, nőknél 26%-át tették ki az irreguláris szaka-

szok a DÜH kifejezésekor. A nők esetében emellett a MEGELÉGEDETTSÉGET kifejező mondatok is nagy, 34%-os arányban tartalmaztak irreguláris zöngét. A leheletes fonáció mind a férfiak, mind a nők esetében a pozitív valenciájú és alacsonyabb aktivációjú érzelmek kifejezése esetén volt nagyobb arányban jellemző, ez nagymértékben egybeesik Laver (1980) megfigyeléseivel, aki a leheletes zöngét az INTIMITÁS és a BIZALMASSÁG kifejezéséhez köti, mely érzelmek szintén alacsony aktivációval és pozitív valenciával jellemezhetők.

A fonációtípusok elemzése mellett elektroglottográf használatával vizsgáltam a hangjak működését a mondatok modális szakaszaiban. Az artikulációs elemzés arra a kérdésre kereste a választ, hogy mérhető-e a modálisnak kategorizált szakaszokon belül artikulációs eltérések a kifejezett érzelem függvényében. Ennek megállapítására a hangjak zárt állapotának a glottális periódus teljes időtartamára levetített arányát (CQ) vizsgáltam. A férfiak adatainál a statisztikai elemzés szignifikáns eltérést mutatott az alacsony és magas aktivációjú érzelmek CQ-értékei között: a férfi adatok között az alacsonyabb CQ-értékekkel jellemezhető, tehát kisebb hangjak-érintkezéssel képzett, leheletesebb zöngét produkálnak az alacsony aktivációjú érzelmek esetén, mint magas aktivációnál. Az aktiváció hatásával interakciót mutatott a valencia: a fent leírt hatás nem, vagy csak kisebb mértékben érvényesült a pozitív valenciájú érzelmek esetében, míg a negatív valenciájú érzelmeknél nagyobb különbségek voltak megfigyelhetők az alacsony és magas aktivációjú érzelmek között. Lehetséges, hogy a nők esetében azért nem volt megfigyelhető ilyen hatás, mivel az általuk produkált zöngé általánosságban is kisebb mértékű hangjak-érintkezéssel jellemezhető (vö. Higgins–Saxman 1991, Ma–Love 2010), és ez az eleve leheletesebb zöngé kisebb változatosságot mutat az érzelmkifejezés során.

A modális szakaszokra vonatkozó artikulációs adatok és a kategorikusan meghatározott fonációtípusok egységes tendenciát mutattak. Az alacsonyabb aktivációjú érzelmek esetében gyakrabban jelent meg leheletes zöngé, és az erre a fonációs típusra jellemző kisebb mediális kompresszió, a hangjak nyitott állapotának nagyobb aránya az alacsony aktivációjú érzelmet kifejező mondatok modális szakaszait is jellemezte. Az érzelmi aktiváció hatása a zöngé minőségére összhangban van a szakirodalomban gyakran megjelenő tendenciával, miszerint a magasabb érzelmi aktiváció nagyobb glottális feszítettséggel jár (vö. Scherer 1986). A kutatás eredményei továbbá alátámasztják azt a feltevést, hogy az artikulációs vizsgálatok az érzelmi aktiváció dimenzióján túl más faktorok, így például a valencia hatását is feltárhatják az emotív kommunikációban: a férfiak esetében a CQ-értékek variációját nem az aktiváció, hanem az aktiváció és a valencia együttesen magyarázzák.

Az alaphérfrekvencia modális szakaszokon mért értékeire szintén hatással volt a kifejezett érzelem: a nemzetközi szakirodalomban leírtakkal összhangban mind a férfiak, mind a nők ejtésében szignifikánsan magasabb alaphérfrekvenciával valósultak meg a magasabb aktivációjú érzelmeket kifejező szakaszok,

mint az alacsonyabb vagy semleges aktivációjúak. A legmagasabb, kiugró alaphangfrekvencia-értékeket a MEGLEPŐDÉS esetén mértem. Az érzelmek valenciája az  $f_0$  értékeire nem volt szignifikáns hatással.

Az adatokon főkomponens-elemzést is végeztem annak vizsgálatára, hogy vajon a dimenzionális érzelmelméleti modellekben gyakran megjelenő dimenziók (aktiváció és valencia) kellőképpen magyarázzák-e az általam mért adatok variabilitását. A kétféle CQ, az  $f_0$  és az érzelm változók varianciáját elemző főkomponens-elemzés által feltárt első két főkomponens nagymértékben egybeesett ezzel a két dimenzióval. A függő változóknál külön-külön megfigyelt hatásoknak megfelelően a főkomponens-analízis alapján is az mondható el, hogy míg az  $f_0$ -értékek szinte kizárólag az aktiváció dimenziójában mutatnak variabilitást, a hangajak-érintkezés aránya a valencia dimenziójától is függ. Az érzelmek összességében inkább az aktivációval összefüggésbe hozható második dimenzió mentén mutatnak eltérést, a valenciával összefüggő első dimenzió mentén nagyobb eltéréseket csak a férfiak csoportján belül tapasztalhatunk.

Az itt bemutatott kutatásban először vizsgáltam együttesen a fonációtípusok, a hangajak-érintkezés és az alaphangfrekvencia variabilitását magyar nyelvű emotív kommunikációban. Tudomásom szerint a szakirodalomban elsőként alkalmaztam főkomponens-elemzést az emotív megnyilatkozások akusztikai és artikulációs paraméterein egy dimenzionális érzelmelméleti modell tesztelésére. További kutatások tárgya lehet más nyelvi egységek vizsgálata, illetve az emotív és az emocionális kommunikáció jellegzetes tendenciáinak összehasonlítása, illetve előrejelző lehet más módszertanok, például inverz szűrés használata is.

### Irodalom

- Awan, Shaheen N. – Awan, Jordan A. 2013. The effect of gender on measures of electroglottographic contact quotient. *Journal of Voice* 27/4. 433–440.
- Bárczi Géza – Országh László (eds.) 1959–1962. A magyar nyelv értelmező szótára I–VII. Akadémiai Kiadó, Budapest.
- Boersma, Paul 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* 17. University of Amsterdam. 97–110.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer*. (v. 6.0.31) <http://www.praat.org/>. (A letöltés ideje: 2017. szeptember 10.).
- Burkhardt, Felix – Paeschke, Astrid – Rolfes, Miriam – Sendlmeier, Walter F. – Weiss, Benjamin 2005. A database of German emotional speech. *Interspeech* 5. 1517–1520.
- Chavent, Marie – Kuentz, Vanessa – Labenne, Amaury – Liquet, Benoit – Saracco, Jerome 2017. *PCAmixdata: Multivariate Analysis of Mixed Data*. R package version 3.1. <https://CRAN.R-project.org/package=PCAmixdata>
- Childers, Donald G. – Krishnamurthy, Ashok K. 1985. A critical review of electroglottography. *Critical reviews in biomedical engineering*. 12/2. 131–161.

- Childers, Donald G. – Lee, Chih K. 1990. Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America* 90/5. 2394–2410.
- Davitz, Joel Robert 1964. *The communication of emotional meaning*. McGraw-Hill, New York.
- Drioli, Carlo – Tisato, Graziano – Cosi, Piero – Tesser, Fabio 2003. Emotions and voice quality: experiments with sinusoidal modeling. *ISCA Tutorial and Research Workshop on Voice Quality: Functions, Analysis and Synthesis*. 127–132.
- Fant, Gunnar 1960. *Acoustic Theory of Speech Production*. Mouton & Co., The Hague, Netherlands.
- Fónagy Iván 1978. A new method of investigating the perception of prosodic features. *Language and Speech* 21/1. 34–49.
- Henrich, Natalie – d'Alessandro, Christophe – Doval, Boris – Castellengo, Michèle 2004. On the use of the derivative of electroglottographic signals for characterization of nonpathological voice phonation. *Journal of the Acoustical Society of America* 115/3. 1321–1332.
- Herbst, Christian 2004. *Evaluation of various methods to calculate the egg contact quotient*. Diploma thesis in music acoustics. KTH Speech, Music and Hearing, Stockholm, Sweden.
- Higgins, Maureen B. – Saxman, John H. 1991. A comparison of selected phonatory behaviors of healthy aged and young adults. *Journal of Speech, Language, and Hearing Research* 34. 1000–1010.
- Hillenbrand, James – Cleveland, Ronald A. – Erickson, Robert L. 1994. Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research* 37/4. 769 – 778.
- Howard, David M. 1995. Variation of Electrolaryngographically Derived Closed Quotient for Trained and Untrained Adult Female Singers. *Log Phon Vocol* 9/2. 163–172.
- Kirby, James. 2017. *Praatdet: Praat-based tools for EGG analysis*. (v. 0.1.1). <https://doi.org/10.5281/zenodo.1117189>.
- Kankare, Elina – Laukkanen, Anne-Maria – Ilomäki, Irma – Miettinen, Anne – Tiina Pyllkkänen 2012. Electroglottographic contact quotient in different phonation types using different amplitude threshold levels. *Logopedics Phoniatrics Vocology* 37/3. 127–132.
- Keating, Patricia – Garellek, Marc – Kreiman, Jody 2015. Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow.
- Kelemen József 1970. *A mondatszók a magyar nyelvben*. Akadémiai Kiadó, Budapest.
- Kugler Nóra 1998. A mondatszók. *Magyar Nyelvőr* 122. 337–347.
- Kuznetsova, Alexandra – Brockhoff, Per B. – Christensen, Rune HB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82. (13). 1–26. <http://doi.org/10.18637/jss.v082.i13>.
- Ladefoged, Peter 1971. *Preliminaries to linguistic phonetics*. University of Chicago, Chicago.
- Laukkanen, Anne-Maria – Vilkmán, Erkki – Alku, Paavo – Oksanen, Hanna 1996. Physical variation related to stress and emotionally state: a preliminary study. *Journal of Phonetics* 24. 313–335.

- Laver, John 1980. *The phonetic description of voice quality*. Cambridge University Press, Cambridge.
- Ma, Estella P. – Love, Amanda L. 2010. Electrolottographic evaluation of age and gender effects during sustained phonation and connected speech. *Journal of Voice* 24/2. 146–152.
- Markó Alexandra 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- Marty, Anton 1908. *Untersuchungen zur allgemeinen grundlegung der grammatik und sprachphilosophie*. Niemeyer, Halle/Saale.
- Michaud, Alexis – Thi, Lan N. 2007. Peakdet. MatLab script. <https://github.com/covarep/covarep/tree/master/glottalsource/egg/peakdet>.
- Menezes, Caroline – Maekawa, Kikuo 2006. Paralinguistic effects on voice quality: a study in Japanese. *Proceeding of Speech Prosody 2006*. 656–659.
- Murray, Iain R. – Arnott, John L. 1993. Towards the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America* 93. 1097–1108.
- Murray, Iain R. – Arnott, John L. 2008. Applying an analysis of acted vocal emotions to improve the simulation of synthetic speech. *Computer Speech & Language* 22/2. 107–129.
- Rosten, Leo 1968. *The joys of Yiddish*. McGraw-Hill, New York.
- Rothenberg, Martin 1992. A Multichannel Electrolottograph. *Journal of Voice* 6/1. 36–43.
- R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>. (A letöltés ideje: 2017. szeptember 10.).
- Russell, James 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39. 1161–1178.
- Scherer, Klaus R. 1986. Vocal affect expression: A review and a model for future research. *Psychological Bulletin* 99. 143–165.
- Sztahó, Dávid 2013. *Automatikus érzelm-felismerés akusztikai paraméterek alapján*. PhD-értekezés, Budapesti Műszaki és Gazdaságtudományi Egyetem, Budapest.
- Stankiewicz, Edward 1964. Problems of emotive language. Sebeok, Thomas A. et al. (eds.). *Approaches to semiotics: Cultural anthropology, education, linguistics, psychiatry, psychology; transactions of the Indiana University Conference on Paralinguistics and Kinesics*. 239–264.
- Webster, Noah 1973. *The international Webster new encyclopaedic dictionary*. Tabor House, New York.
- Wierzbicka, Anna 1992. The semantics of interjection. *Journal of pragmatics* 18/2–3. 159–192.
- Zagalo, Nelson – Torres, Ana – Branco, Vasco 2005. Emotional spectrum developed by virtual storytelling. *International Conference on Virtual Storytelling*. Springer Berlin Heidelberg. 105–114.

### Köszönetnyilvánítás

A kutatás a Nemzet Fiatal Tehetségeiért Ösztöndíj (NTP-NFTÖ-17-B-0342) keretein belül valósult meg. Köszönöm Markó Alexandrának a kutatás során nyújtott áldozatos

témavezetői segítséget, Deme Andreának az előremutató javaslatokat, valamint Siptár Péternek és Gocsál Ákosnak az OTDK-bírálat formájában megfogalmazott észrevételeiket.

### **Phonatory variation in acted emotive speech**

The aim of this paper is to describe how emotive information is conveyed through speech. Specific attention was paid to the differences along the dimension of emotional arousal and valence. 10 Hungarian speakers, 5 female and 5 male uttered the vowel /a/ expressing 7 emotions and an emotionally neutral realization, 3 repetitions each. Speech and electroglottographic (EGG) signals were recorded simultaneously in a sound-treated room. Voice quality was determined based on both auditory impressions and visual inspection of waveforms and spectrograms. For periods labeled as modal, fundamental frequency ( $f_0$ ) was measured, and the EGG contact quotient (CQ) was calculated using two different evaluation techniques for a fine-grained articulatory analysis of this perceptually homogenous and neutral phonation type. Apart from analyzing how emotions showing different activation and valence differ from each other with respect to these parameters using generalized linear mixed models, a principal component analysis (PCA) of the above mentioned acoustic and articulatory measures was also carried out to inspect whether variability of these parameters is, in fact, best explained by these two dimensions. Irregular phonation was found to be similarly frequent in the expression of ANGER for both genders (26% of the total duration of all utterances expressing that emotion for females, 29% for males). Females also used irregular phonation often when expressing CONTENTMENT (34%), a pattern not appearing in males' speech. Breathy voice was observed to express emotions with positive valence (CALM and HAPPINESS) for both genders. Males followed this pattern even in the case of CONTENTMENT (30%), showing that this emotion is expressed using different non-modal voice qualities. Fundamental frequency was significantly higher when expressing emotions with high arousal, compared to low-arousal emotions. CQ values did not show any systematic variation for females, but for males, there was a significant effect of emotion on CQ. Particularly, the effect of arousal level was significant: a Tukey post-hoc test showed that men express low-arousal emotions with significantly lower CQ compared to emotions with high arousal. The principal component analysis of  $f_0$  and the two CQ measures showed that the variability of these measures is best explained by dimensions that actually cluster emotions of similar arousal and valence together.

## EGYES TESTALKATI PARAMÉTEREK ÉS A KITÖLTÖTT SZÜNETEK FORMÁNSSZERKEZETE KÖZÖTTI ÖSSZEFÜGGÉS VIZSGÁLATA FÉRFI BESZÉLŐKNÉL

Gocsál Ákos

PTE MK, MTA Nyelvtudományi Intézet

### Bevezetés

A legtöbbször megtapasztalhattuk – írja Krauss (2002) egyik munkájában –, hogy amikor először látunk egy olyan személyt, akit előtte csak a hangja alapján, például rádióból vagy telefonon keresztül ismertünk, akkor igencsak meglepődünk az illető kinézetén. Ennek oka az, hogy a hangokhoz sztereotípiákat rendelünk, és ha a vizuális csatornán nem áll rendelkezésünkre információ, akkor sztereotípiáink alapján elképzeljük a beszélőt. Ezek a sztereotípiák ugyanakkor gyakran félrevezetőek, a benyomás kialakításához felhasznált, észlelt beszédparaméterek a valóságban sokszor egyáltalán nem korrelálnak a beszélő tényleges tulajdonságaival.

Jelen munkánk központi kérdése ennek a problémakörnek egy részterülete: egyes testalkati paraméterek – nagyobb hangsúllyal a testmagasság, de emellett a testsúly és a testtömegindex – illetve a beszéd akusztikai struktúrái közötti esetleges összefüggések vizsgálata. Vissza lehet-e következtetni az akusztikai paraméterekből a beszélő testmagasságára? Személyészlelési kísérletekben számos kutatás igazolta néhány beszédakusztikai paraméter sztereotipikus szerepét. Ismeretes például, hogy a mélyebb alaphang magasabb, testesebb beszélő benyomását kelti a hallgatókban (van Dommelen–Moxness 1995; Feinberg et al. 2005; Pisanski–Rendall 2011). Az alaphang és a testmagasság között feltételezett kapcsolat objektív vizsgálata azonban nem igazolt korrelációt (Künzel 1989; Evans et al. 2006; González 2007). Ennek okát Fitch (1997) abban látta, hogy a gégeporcok a test többi részétől függetlenül növekednek, így a gége, és különösen a hangszalagok mérete a többi testrésztől függetlenül alakulhat, különösen a férfiaknál, akiknél serdülőkorban a gége, a hangszalagok, és a toldalékcső növekedése a magas tesztoszteronszint miatt gyorsabb és erőteljesebb (Puts et al. 2007). Emellett pedig az alaphang használata kultúrafüggő is lehet (Evans et al. 2006). A toldalékcső hossza és a beszélő testmagassága között azonban igen szoros ( $r = 0,926$ ;  $p < 0,0001$ ) korrelációt állapított meg Fitch és Giedd (1999), így kézenfekvőnek tűnik az a feltételezés, hogy a toldalékcső hosszával összefüggő beszédakusz-

Gocsál Ákos 2018. Egyes testalkati paraméterek és a kitöltött szünetek formánsszerkezete közötti összefüggés vizsgálata férfi beszélőknél. *Beszédkutatás* 2018. 63–84.



tikai paraméterekből megbízhatóan lehet visszakövetkeztetni a beszélő testmagasságára.

Ennek matematikai leírásához kiindulópontként jól használható Fant ismert forrás-szűrő elmélete, amelyet a szerző jelen munkánk szempontjából fontos szempontokkal is kiegészített (pl. Fant 1966, 1967). A legegyszerűbb akusztikai modell szerint a toldalékcső egyik végén zárt, másik végén nyitott lineáris cső, amelyben a negyed, háromnegyed, öt negyed stb. hullámhossznál jelentkeznek a rezonanciafrekvenciák, amelyek az (1) számú képlet segítségével határozhatók meg (pl. Fitch 1997; Johnson 2003:96). A képletben  $i$  jelzi az adott formáns sorszámát,  $c$  a hang terjedési sebességét,  $L$  a cső hosszát.

$$F_i = \frac{(2i - 1)c}{4L} \quad (1)$$

A képlet alapján kiszámolható, hogy hol várhatóak a formánsfrekvenciák, ha ismert a toldalékcső hossza. Egy férfi beszélőnél átlagosnak tekinthető, 17,5 cm-es toldalékcső esetében,  $c = 340$  m/s hangsebességgel számolva, az első formánst megközelítőleg 500, a másodikat 1500, a harmadikat 2500 Hz közelében várhatjuk. Megfordítva, a formánsfrekvenciák ismeretében, az egyenlet átrendezésével következtetni lehet a toldalékcső hosszára (van Dommelen–Moxness 1995; Fitch–Reby 2001; Dusan 2005).

$$L = \frac{(2i - 1)c}{4F_i} \quad (2)$$

Ennek alapján pedig – mivel a fentiekben már láttuk, hogy a toldalékcső hossza és a testmagasság erősen korrelál –, várhatjuk, hogy a testmagasság jó közelítéssel megbecsülhető.

Az (1) és (2) számú képletek azonban csak abban az esetben közelítik jól a valós értékeket – akár formánsfrekvenciákat, akár a toldalékcső hosszát kívánjuk kiszámítani –, ha a toldalékcső alakja ténylegesen hasonlít a félig nyitott csőhöz, nincsenek benne jelentős szűkületek. Ez a semleges magánhangzó, az [ə] esetében érvényesül (Hayward 2000: 88), jól artikulált magánhangzók esetében azonban már kevésbé. Ennek oka az, hogy míg az [ə] képzésekor a cső átmérője jó közelítéssel egyenletesnek tekinthető (van Dommelen–Moxness 1995), addig a jól artikulált magánhangzók képzésekor szűkületek keletkeznek, amelyek a képzett magánhangzó jellegének megfelelően befolyásolják a cső rezonancia-válaszgörbáját. Ennek a problémának a legegyszerűbb feloldása az, ha a toldalékcsővet két, egymáshoz illesztett, különböző átmérőjű csőnek tekintjük. A „hátsó” cső, gége és a nyelvhat pozíciójából adódó szűkület közötti, garati üreg akusztikai szempontból zárt csőnek tekinthető, míg a nyelvhat és a szájnyílás közötti, „elülső”, szájüregi rész félig nyitott csőként modellezhető. Fant (1966) korábbi, forrás-szűrő elméletét, illetve ezeket a megfontolásokat alapul véve közölt formánseredményeket. Radiológiai vizsgálattal megállapította egy beszélő toldalékcsővének hosszát,

majd megmérte, illetve az általa felállított modell és az anatómiai adatok alapján kiszámította az [i] hang formánsait. A legnagyobb egyezést F2 esetében találta, a férfi beszélőnél a két érték megegyezett (2060 Hz), a női beszélőnél mindössze 50 Hz volt a különbség (mért: 2600 Hz, számított: 2650 Hz). Az 1. formáns két értékére a női beszélőnél szintén kicsi (mért: 240 Hz, számított: 247 Hz), a férfi beszélőnél nagyobb (mért: 225 Hz, számított: 156 Hz) különbséget kapott. Hasonló modell alapján közölt formánsértékeket Lindblom és Sundberg (1972), akik ugyanakkor felvetették – szintén Fant 1960-ban publikált, eredeti modelljére hivatkozva –, hogy a szűkületnek is van egy meghatározható hosszúsága, tehát még pontosabb eredmények kaphatók, ha a toldalékcso teljes hosszát három különböző átmérőjű, azaz garati, szűkületi és szájúregi szakaszra osztják fel.

Megjegyzendő, hogy léteznek más, bonyolultabb matematikai apparátust alkalmazó, több szakaszból építkező modellek is. Ahmadi és McLoughlin (2012) például – mások által, korábban leírtak alapján – 22, különböző átmérőjű szakaszból épített fel egy mesterséges toldalékcsovet, amely igen jó minőségben produkált magánhangzókat. Ennek ellenére, a két, illetve három csőszakasz alapján építkező modellek is széles körben elterjedtek, ugyanis már ezek is jól magyarázzák a magánhangzó-minőségeket, ugyanakkor jóval egyszerűbb a matematikai apparátusuk.

A toldalékcso méretének becslésével kapcsolatban több szerző munkájában is megjelenik egy Fitch (1997) által bevezetett paraméter, a formánsdiszperzió (pl. Rendall et al. 2005; Evans et al. 2006). Ez a paraméter azt mutatja, hogy a szomszédos formánsok átlagosan milyen távolságra helyezkednek el egymástól. Kiszámítása a (3) képlet alapján történik, amelyben  $N$  az összes vizsgált formáns száma,  $i$  az aktuálisan vizsgált formáns sorszáma,  $F$  a frekvenciája.

$$D_f = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N-1} \quad (3)$$

Ebből levezetve, az első négy formánsra vonatkozóan a formánsdiszperzió az alábbiak szerint számítható ki (Rendall et al. 2005):

$$D_f = \frac{(F_2 - F_1) + (F_3 - F_2) + (F_4 - F_3)}{3} \quad (4)$$

Fitch (1997) azt feltételezte, hogy az így kiszámított diszperziós érték is tükrözi a testméretet. Rendall és munkatársai (2005), elismerve, hogy egyre többen alkalmazzák ezt a paramétert, kritikai véleményt is megfogalmaztak. Azt állapították meg, hogy mivel az F1 variabilitása a legkisebb, az F4-é pedig a legnagyobb, az említett (3) jelű számítási mód lényegében az F4 és az F1 különbségére redukálja a formánsdiszperzió értékét. Ez pedig akár azt is jelentheti, hogy a formánsdiszperziót csak az F4 határozza meg. Ennek a problémának egy lehetséges feloldása a formánsarányok kiszámítása. Smith

és munkatársai (2007) az F2/F1 és F3/F1 értékeket határozták meg és vontak le belőlük következtetéseket.

Bár a fenti – és egyéb, itt nem közölt – formulák alapján azt várnánk, hogy a formánsstruktúra jól tükrözi a testmagasságot, a kísérleti eredmények között sok a bizonytalanság és ellentmondás. Nem egyértelmű például, hogy melyik formáns, és milyen módon függ össze a testmagassággal. Pisanski és munkatársai (2014a) felvetik, hogy egyes formánsok megbízhatóbban jelezhetik a testméreteket, mint más formánsok, de bizonytalan, hogy melyek azok. Részben ugyanők egy másik munkájukban arra a következtetésre jutottak, hogy az első négy formáns közül a magasabbak megbízhatóbban tükrözik a testméreteket, mivel az F1-nek és az F2-nek elsősorban a beszédértésben van szerepük (Pisanski et al. 2014b). Hasonló véleményt fogalmazott meg Barreda (2016), aki ezt azzal magyarázta, hogy a mélyebb formánsok elsősorban fonémaspecifikus információkat hordoznak. Emiatt igen nagy lehet a változatosságuk, amelyek mellett a toldalékcsohosszak közötti, viszonylag kis különbségek eltűnhetnek. Javaslatuk szerint a testmagasságok becslésekor inkább az F3-t és az F4-et kellene alapul venni.

González (2004) az alaphang és öt különböző, izoláltan ejtett magánhangzó első négy formánsa és a beszélő testmagassága közötti korrelációkat vizsgálta. A férfi beszélők esetében egyedül az /e/ hang második formánsa, a női beszélőknél az /a/, /e/, /i/ hangok második, illetve az /e/ és /i/ hangok harmadik formánsa korrelált a testmagassággal. A korrelációs együttható minden esetben negatív előjelű volt, azaz a magasabb beszélőknél mélyebb formánsfrekvenciák adódtak, de a szerző, más mérései alapján is, igen gyenge összefüggésre következtetett mérési adatai alapján. Hatano és munkatársai (2012) az előzőekhez részben hasonló, részben azoktól eltérő eredményeket kapott. A férfi beszélők által izoláltan ejtett /a/, /e/, /i/, /o/ és /u/ magánhangzók realizációi közül csak az /e/ első, második és negyedik formánsa korrelált a toldalékcsohosszával, szintén negatív előjellel. A szerzők, ellentétben korábbi eredményekkel (Fitch–Giedd 1999), semmilyen korrelációt nem találtak az MRI-felvételek segítségével megállapított toldalékcsohosszak és a formánsfrekvenciák között, és a testmagasság sem korrelált a formánsfrekvenciákkal. Ugyanakkor gyenge, negatív előjelű összefüggés adódott az  $f_0$  és a testmagasság között, ami szintén arra utal, hogy az alacsonyabb beszélők magasabb alaphangon beszéltek, és fordítva, ami szintén ellent mond korábbi adatoknak, amelyek a testmagasság és az alaphang függetlenségét mutatták.

Dusan (2005) a TIMIT amerikai beszédkorpuszból beszélőnként 10 mondatot használt fel, amelyekből 20-20 magánhangzó-variánst emelt ki (pl.: /iy/, /ih/, /eh/ stb.). Több akusztikai paraméter kombinációjával (Mel-frekvenciás kepsztrum, LPC-k, formánsfrekvenciák, alaphang) 57%-os magyarázóerejű ( $R = 0,775$ ) lineáris regressziós modellt kapott, de az alaphang, illetve a formánsfrekvenciák és a testmagasság között külön elvégzett korrelációs számítások is szignifikáns kapcsolatot mutattak ki. A szerző 462 beszélő adatait fel-

használva, magas korrelációs együtthatót kapott ( $r = 0,721$ ) a toldalékcso becsült hossza és a beszélő testmagassága közötti összefüggésre.

Rendall és munkatársai (2005) izolált magánhangzókon, hVd és bVt szerkezetű szavakban ejtett magánhangzókon, egy szótagos szavakból álló mondatok magánhangzóin végeztek méréseket. A női beszélők esetében sem az alaphang, sem a formánsdiszperzió, de az első négy formáns sem adott szignifikáns magyarázóerőt, amikor a regressziós modellben a beszélő testmagasságát tekintették függő változónak. A férfiak esetében azonban a formánsdiszperzió szignifikáns magyarázóerőt adott a regressziós modellben a magánhangzóknál és az [ə] hangoknál is. A magánhangzóknál a 4. formáns  $p < 0,01$  szinten, az [ə] hang esetében pedig a 4. formáns  $p < 0,01$  szinten, az  $f_0$ , az F1 és az F2 pedig  $p < 0,05$  szinten adott szignifikáns magyarázóerőt, amikor a függő változó a beszélő testmagassága volt. Egy másik kutatásban viszont a formánsdiszperzió negatív előjellel, szignifikánsan korrelált a testmagassággal, ami azt jelentette, hogy magasabb beszélők esetében kisebb diszperziós értékek adódtak (Evans et al. 2006), azaz, a formánsok a magasabb beszélőknél sűrűbben jelentek meg.

Smith és munkatársai (2007) a korábban említett formánsarányokat vizsgálták meg hosszan kitartott magánhangzókon. A 170 cm-nél magasabb beszélőknél az arányszámok magasabbak voltak, mint a 150 cm alatti beszélőknél. A legnagyobb különbség az F3/F1 aránynál mutatkozott. Az eredmények azonban óvatosan kezelendők, mivel csak négy (két magas és két alacsony, gyermekkorú) beszélővel végezték a kísérletet.

Megemlítendő még, hogy egyes szerzők a testmagasság és a testtömeg négyzetének hányadosaként meghatározott testtömegindexet (*Body Mass Index*, BMI) használják az akusztikai paraméterekkel való összehasonlításhoz. Egy női beszélőkkel végzett kutatásban a BMI alapján túlsúlyosnak, illetve elhízottnak minősülő személyek alaphangja szignifikánsan mélyebb volt, mint azoké, akiknél a BMI a normál tartományba esett. Nem szignifikáns mértékben, de valamivel alacsonyabb volt az alaphang azoknál is, akiknél a normálnál kisebb BMI-t állapítottak meg (Souza–Santos 2017). Férfiaknál Evans és munkatársai (2006) azt találták, hogy az alaphang nem korrelált a BMI-vel, a testsúly azonban igen, a nagyobb testsúly szignifikánsan mélyebb alaphanggal járt együtt. Hadman és munkatársai (2013) pedig a BMI és két magánhangzó első négy formánsa között kerestek összefüggést, de nem találtak szignifikáns kapcsolatot.

További fontos kérdés annak vizsgálata, hogy a beszélő testmagasságának hang alapján történő megbecslése mennyire pontos, illetve a lehallgatók felhasználják-e a formánsstruktúrát a beszélő testmagasságának becsléséhez.

Gianni és munkatársai (1989) kimutatták, hogy a lehallgatók a 169 cm körüli testmagasságú férfi beszélő magasságát általában jól ítélik meg, az ennél magasabb beszélőket azonban jellemzően alul-, a 169 cm-él alacsonyabb beszélőket pedig felülbecslik. Hasonlóan alakultak az eredmények a női beszé-

lőknél is, azzal a különbséggel, hogy – az általuk közölt grafikon alapján – a kb. 163 cm magasságú beszélőkre jósolt a modelljük helyes becslést, míg az ennél magasabb beszélőket jellemzően a valósnál jellemzően alacsonyabbnak, a 163 cm-nél alacsonyabb beszélőket pedig magasabbnak vélték a lehallgatók. Mindkét esetben megfigyelhető volt, hogy minél távolabb esett a beszélő valós testmagassága a helyes becslés helyétől (azaz a 169, ill. 163 cm-től), annál nagyobb volt az alul- illetve felülbecslés mértéke.

Régebbi kutatások még arról számolnak be, hogy a lehallgatók helytelen sztereotípiaként a beszélő alaphangját használják a testméretek megbecsléséhez – ezt az álláspont ma sem vitatható –, ugyanakkor ugyanilyen félrevezető, sztereotípián alapuló percepciók kulcsnak vélték a formánsszerkezetet is (González 2003).

Charlton és munkatársai (2013) 10 formánst tartalmazó, szintetizált hangokkal végzett kísérleteket. A formánsok az emberi toldalékcso méretének megfeleltek, továbbá a kísérlethez használt alaphangok is az ember által képezhető tartományba estek. A formánsszerkezetükben kismértékben különböző hangpárok közül a férfi lehallgatók szignifikánsan nagyobb arányban állapították meg helyesen, hogy melyik tartozhatott a „nagyobb” beszélőhöz. A testmagasságok becslése a mélyebb alaphangok esetében pontosabb volt, amit a szerzők annak tulajdonítottak, hogy ezekben az esetekben a sűrűbben elhelyezkedő harmonikusoknak köszönhetően a formánsok is egyértelműbben hallhatók.

Feinberg et al. (2005) magánhangzókat játszottak le kísérleti személyeknek. A kísérletek előtt a magánhangzókat többféleképpen is manipulálták. Egyrészt, létrehoztak olyan változatokat, amelyekben a formánsstruktúra változatlanul maradt, az alaphang viszont az eredetihez képest 20 Hz-cel mélyebb, illetve magasabb volt. Másrészt, a formánsstruktúrát változtatták meg úgy, hogy az formánsdiszperzió az eredetinek 95 illetve 105 %-a legyen. Ez utóbbival a különböző toldalékcso-hosszakat szimulálták. A lehallgatók „nagyobb” ítélték az elképzelt beszélőt, amikor a hallott magánhangzó mélyebb alaphanggal szólalt meg, illetve olyan formánsstruktúrával, amelyet hosszabb toldalékcsoval képzett. Hasonlóképpen, a hosszabb toldalékcsoval képzett magánhangzók alapján egy hétfokú skálán magasabb beszélőre következtettek a lehallgatók Smith és Patterson (2005) kísérletében. Ezt a szerzők a toldalékcso hossza és a testméretek közötti erős korrelációval magyarázták.

A hazai szakirodalomban Gósy (2001) foglalkozott hasonló kérdésekkel. Kísérletei azt mutatták, hogy a lehallgatók a mély hangú nők magasságát 52%-ban ítélték meg helyesen, míg a magas hangúakét csak 37,5%-ban. A mély hangú férfiak magasságát pedig 70,8%-ban, a magasabb hangúakét 50%-ban becsülték meg helyesen.

Jelen munkánkban magyar anyanyelvű beszélők hangmintáin végzünk méréseket azzal a céllal, hogy a testmagasság és akusztikai paraméterek – alap-

hang és formánsszerkezet – között összefüggéseket keressünk. A korábbi kutatási eredményeket jellemzően kitartott, izolált magánhangzókra, esetleg szintén izoláltan ejtett szavakra vagy szótagokra vonatkozóan közölték a probléma kutatói. Ugyanakkor felmerül a kérdés, hogy a kimutatott összefüggések, a megfogalmazott bizonytalanságok spontán beszéden mért adatok esetében is ugyanúgy jelentkeznek-e. A spontán beszéden történő mérések azért fontosak, mert egyes alkalmazásoknál, például kriminalisztikai feladatoknál kutatói instrukciók alapján ejtett szavak, izolált hangzók nem állnak rendelkezésre.

A spontán beszéd egyik gyakori eleme a kitöltött szünet, amelynek pszicholingvisztikai háttéréről és akusztikai-fonetikai sajátosságairól már magyar anyanyelvű beszélőkre vonatkozóan is részletes adatok állnak rendelkezésre (pl. Bóna–Vakula 2013, Gósy et al. 2013, Gyarmathy 2015, Laczkó 2013). Az egyéni sajátosságok közül a korábbi kutatások az életkor lehetséges differenciáló szerepét vizsgálták. Azzal kapcsolatban azonban nem állnak rendelkezésre adatok, hogy esetleg más egyéni tulajdonságok, például a beszélő testalkatára jellemző paraméterek tükröződnek-e a kitöltött szünetek valamely akusztikai-fonetikai paraméterében. Jelen kutatásunk e probléma egy részterületének vizsgálja az alábbi kutatási kérdések alapján:

1. Kimutatható-e korreláció testmagasság és az alaphang, illetve formánsstruktúra között? Igazolható-e, hogy valamelyik formáns korrelációja erősebb a testmagassággal, mint más formánsoké?
2. A különböző életkorú személyeknél találhatók-e eltérések a kitöltött szünetek formánsszerkezetében?
3. Létezik-e összefüggés a formánsdiszperzió, illetve a formánsarányok és a testmagasság között?
4. Kimutatható-e összefüggés a mért beszédakusztikai paraméterek és a testtömegindex, esetleg a testsúly között?
5. A formánsértékek alapján megbecsült toldalékcsohosszak korrelálnak-e a beszélők valós testmagasságával?

A kérdésekhez kapcsolódóan hipotéziseket fogalmazunk meg:

1. Feltételezzük, hogy az alaphang nem, de a formánsstruktúra korrelál a testmagassággal. A magasabb beszélőknél mélyebb formánsértékeket várunk.
2. A különböző életkorú személyeknél eltérő formánsszerkezet mutatkozik.
3. A formánsarányok és a formánsdiszperzió különböznek a különböző testmagasságú beszélőknél. A magasabb beszélőknél feltételezett mélyebb formánsértékből kisebb arányok és diszperziós értékek adódnak.
4. A nagyobb testsúlyú, testtömegindexű személyek kitöltött szüneteiben mélyebb alaphangot, mélyebb formánsfrekvenciákat találunk.
5. A formánsértékek alapján kiszámított toldalékcsohosszak és a valós testmagasságok korrelálnak egymással.

**Kísérleti személyek, anyag, módszer**

A kísérletben ugyanannak a 24 férfi beszélőnek a beszédfelvételeit használtuk, akik korábbi kutatásunkban is szerepeltek (Gocsál 2017). A hangfelvételek a BEA adatbázisból származnak (Gósy et al. 2012), a beszélők életkora 20–72 év. Életkoruk eloszlása megközelítőleg egyenletes volt az említett tartományban. Két, életkorukat tekintve egymást követő beszélő között átlagosan 2,16 év volt a korkülönbség, ténylegesen 0–4 éves lépésköz volt közöttük. A beszélők testmagassága a 163–197 cm közötti tartományba esett, átlagos testmagasságuk 179,8 cm (SD=8,9 cm) volt, a testmagasságok normál eloszlást mutattak, kiugró érték nem volt. Testsúlyuk 60–100 kg, átlagosan 82 kg (SD=10,8 cm), az értékek szintén normál eloszlást mutattak, kiugró érték nélkül. A testtömegindex értéke 18,6 és 31,56 közé esett, átlagértéke 25,45 (SD=3,64), kiugró érték nem volt. A beszélők felsőfokú végzettségűek vagy egyetemi hallgatók, és nem dohányzók voltak.

Az akusztikai elemzéshez a BEA protokoll szerinti „interjú”, „érvelés”, illetve „társalgás” című részekből válogattunk *ő*-féle kitöltött szüneteket. Eredeti szándékunk szerint beszélőnként megközelítőleg azonos számú, olyan kitöltött szüneteket kívántunk felhasználni, amelyek előtt és után csendes szünetek voltak, hogy a koartikulációs hatásokat, amennyire csak lehet, kiküszöböljük. Ilyen kitöltött szünet azonban viszonylag kevés volt, ezért a méréseket olyan kitöltött szüneteken is elvégeztük, amelyek más környezetben, jellemzően szavak elejéhez vagy végéhez tapadva fordultak elő. A változatos előfordulások miatt beszélőnként jellemzően – de nem minden esetben – 20 kitöltött szüneten végeztük el a méréseket, majd a kapott értékeket átlagoltuk. A mérések során azt találtuk, hogy az egyik beszélő még így is rendkívül kevés *ő*-féle kitöltött szünetet produkált, emiatt a nála mért értékeket a további számítások során nem vettük figyelembe, azaz összesen 23 beszélő adatait használtuk fel. A mérésekhez összesen 444 kitöltött szünetet dolgoztunk fel. 16 beszélőnél 20, két beszélőnél 15, három beszélőnél 22, egy beszélőnél 28 kitöltött szünetet jelöltünk ki a kísérlethez. Ez utóbbi esetekben azért volt szükség 20-nál több hangra, mert az adott beszélők beszédében nagy arányban voltak irregulárisan képzett zöngék, és a hangzók közepén, ahol a méréseket végeztük, több esetben további számításokra nem alkalmas  $f_0$ -értékek adódtak. Két esetben pedig csak 15 kitöltött szünetet találtunk a rendelkezésre álló beszédmintákban, de ezt már elegendőnek ítéltük a számításokhoz.

Az akusztikai méréseket a Praat program segítségével végeztük (Boersma–Weenik 2014). A mérésre kijelölt kitöltött szünetek közepén az alaphangot, illetve az első öt formánst mértük meg automatikusan, egy erre a célra írt script segítségével. Babel és munkatársai (2014) közlése szerint az F5 nem mérhető megbízhatóan, amit saját adataink is részben megerősítettek. Három beszélőnél valóban, viszonylag kevés – 5, 7, illetve 9 – kitöltött szünet esetében volt az F5 mérhető, a többi esetben azonban legalább 10–15 F5-értéket kaptunk. A script futtatását követően az adatsorokat minden egyes beszélő

esetében ellenőriztük, és ahol szükséges volt, ott adattisztítást, illetve -korrekciót hajtottunk végre. Egyrészt, az irregulárisan képzett zöngékhez tartozó, extrém mély  $f_0$ -értékeket eltávolítottuk (vö. Markó 2013:19) és csak a modális zöngéken állapítottuk meg az átlagos alaphangot. Másrészt, az automatikusan kinyert formánsadatok több helyen pontatlannak bizonyultak, különösen a legmélyebb hangú, illetve a rekedtesen vagy irreguláris zöngékkel beszélőknél. Gyakori hiba volt, hogy az F1-et nem észlelte a script, és az F2 értékét sorolta F1-nek, illetve az algoritmus nem talált helyes formánsértékeket, annak ellenére, hogy a formánsok a spektrogramon jól kirajzolódódtak. Ezekben az esetekben a formánsértékeket a Praat programban manuálisan, a spektrogram segítségével, a szátkereszt megfelelő pozicionálásával állapítottuk meg.

Az összefüggések igazolásához a Pearson-féle korrelációs együtthatót számítottuk ki az SPSS 23 program használatával. Emellett lineáris regressziós modelleket állítottunk fel, melyekben függő változóként a testalkati paraméterek, magyarázó változóként pedig a mért, illetve kiszámított beszédakusztikai paraméterek szerepeltek. Minden esetben  $p < 0,05$  esetében tekintettük szignifikánsnak a vizsgált kapcsolatot.

### Eredmények

Az 1. táblázat tartalmazza a fenti módszerekkel megállapított átlagértékeket és szóródásokat az összes beszélőre vonatkozóan, összesítve. Az adatok megerősítik, hogy a kutatáshoz felhasznált kitöltött szünetek valóban  $\emptyset$ -féle jellegűek. Gósy (2004: 116) adatai alapján, a köznyelvi magánhangzókkal összehasonlítva a formánsadatok az  $[\emptyset]$ , illetve az  $[\emptyset:]$  formánsaihoz állnak legközelebb (a hivatkozott munkában az  $[\emptyset]$  1. és 2. formánsa: 380 és 1600 Hz, az  $[\emptyset:]$ -é 400 és 1550 Hz). Megállapítható az is, hogy az F1- és F2-értékek is a Gósy et al. (2013) által, a  $[\emptyset]$  hangokra közt 320–640 Hz, illetve 1300–1900 Hz közötti tartományba esnek. A legnagyobb hasonlóságot Horváth (2014:40) eredményeivel találtuk, adatai szerint a férfi beszélők ejtésében vizsgált  $[\emptyset]$  hangok 1. formánsának 447,2 Hz, 2. formánsának pedig 1543 Hz volt az átlagértéke.

1. táblázat: A mért alaphang- és formánsadatok (Hz-ben megadva)

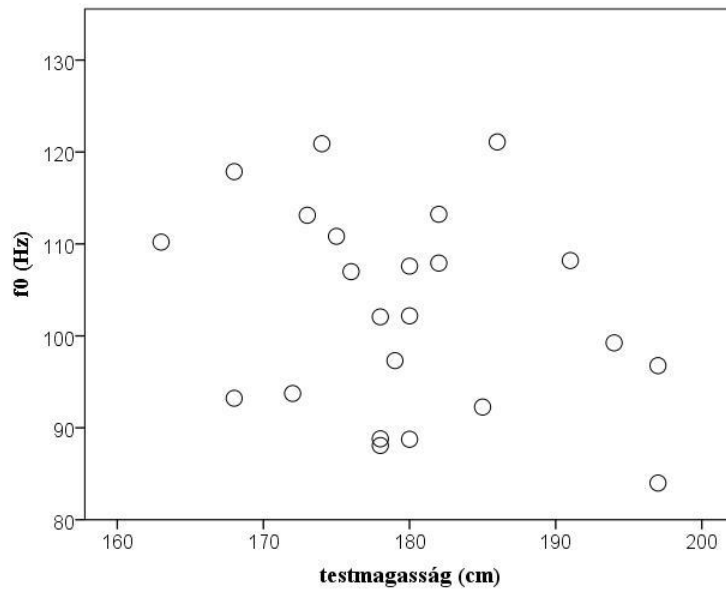
	$f_0$	F1	F2	F3	F4	F5
<b>átlag</b>	102,49	484,60	1587,08	2505,22	3637,57	4707,49
<b>SD</b>	11,48	78,71	177,49	182,22	231,56	241,92

Megvizsgáltuk a formánsok egymás közti korrelációit is. Az F1 és az F3 egyetlen más formánssal sem korrelált, még tendenciaszerűen sem. Az F2 és az F4 között  $r = 0,474$  korrelációs együtthatóval,  $p < 0,05$  szinten, az F4 és az



F5 között pedig  $r = 0,584$  korrelációs együtthatóval,  $p < 0,01$  szinten találtunk korrelációt.

Az alaphang és a testmagasságok között az 1. ábrán látható összefüggést kaptuk. A korrelációs számítás eredménye:  $r = -0,287$ ;  $p = 0,184$ , ami az alaphang és a testmagasság függetlenségét mutatja.



1. ábra

A kitöltött szüneteken mért alaphang és a testmagasság összefüggése

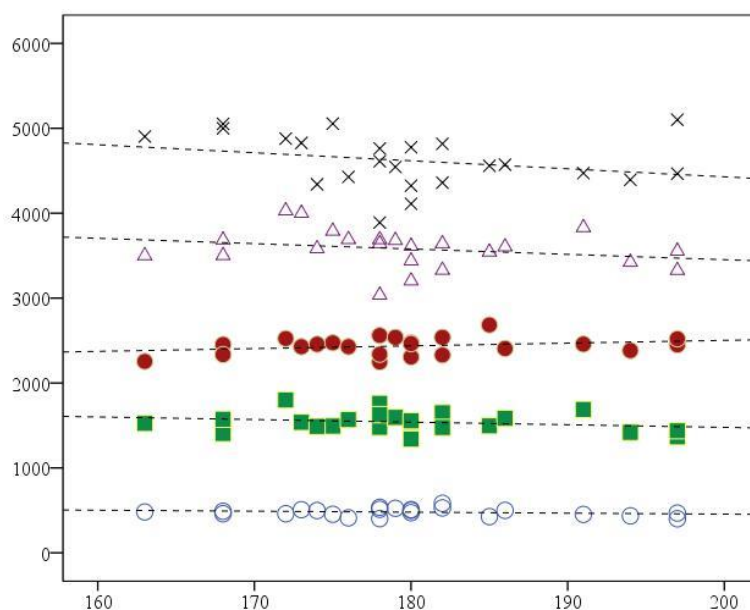
A következő lépésben összefüggéseket kerestünk a formánsértékek és a beszélő testmagassága között. A 2. táblázat tartalmazza a korrelációs együtthatókat, a 2. ábra pedig az adatokat grafikusán ábrázolja. Szignifikáns kapcsolatot egyetlen esetben sem találtunk.

2. táblázat: A testmagasság és formánsok korrelációs együtthatói

	F1	F2	F3	F4	F5
r	-0,227	-0,236	0,276	-0,244	-0,267

Az adatok alapján az ENTER módszerrel lineáris regressziós modellt is megpróbáltunk felállítani, de a 2. táblázatban közölt adatokkal összhangban

nem adódott szignifikáns eredmény ( $F(5,17) = 1,531$ ; nem szignifikáns, a továbbiakban: n.sz.).



2. ábra

A kitöltött szünetek formánsértékei a testmagasság függvényében

Második kérdésünk megválaszolásához megvizsgáltuk, hogy a kitöltött szünetek formánsszerkezete függ-e a beszélő életkorától. Korábbi kutatások ugyanis egyes magánhangzók esetében találtak ilyen jellegű összefüggéseket (Gósy-Bóna 2014). A 3. ábrán látható pontdiagram ábrázolja a kapott adatokat.

A korrelációs számítás ebben az esetben sem adott szignifikáns eredményt. A 3. táblázat tartalmazza a korrelációs együtthatókat, a 3. ábra pedig grafikusán ábrázolja a kapott adatokat.

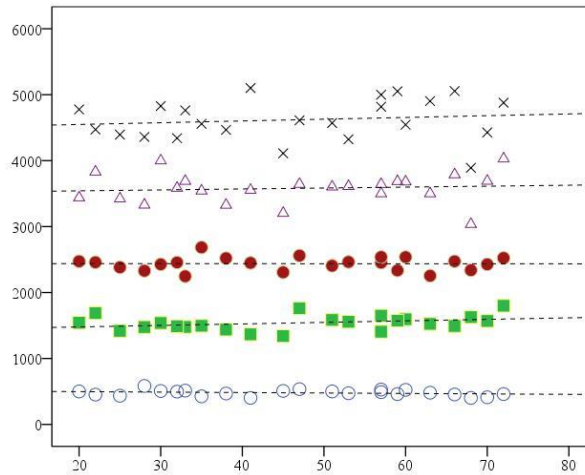
Egy újabb lineáris regressziós modellt kíséreltünk meg felállítani a kapott adatokkal, de ebben az esetben sem adódott szignifikáns magyarázóerejű modell ( $F(5,17) = 1,210$ ; n.sz.).

Harmadik kutatási kérdésünk a formánsdiszperzió és a testmagasság összefüggéseire vonatkozott. Ehhez kiszámítottuk az F2 és az F1 különbségét ( $D_{1-2}$ ), illetve a (3) jelű képlet alkalmazásával formánsdiszperzió értékeit a 3., 4., és 5. formánsig bezárólag ( $D_{1-3}$ ,  $D_{1-4}$ ,  $D_{1-5}$ ), a, minden egyes beszélőre vonat-

kozóan. A 4. táblázatban szereplő korrelációs együtthatók egyike sem utal arra, hogy a formánsdiszperzió értékei korrelálnának a beszélők testmagasságával.

3. táblázat: Az életkor és a formánsok korrelációs együtthatói

	<b>F1</b>	<b>F2</b>	<b>F3</b>	<b>F4</b>	<b>F5</b>
<b>r</b>	-0,245	0,328	-0,009	0,101	0,138



3. ábra

A kitöltött szünetek formánsai az életkor függvényében

4. táblázat: A testmagasság és a formánsdiszperziós értékek korrelációs együtthatói

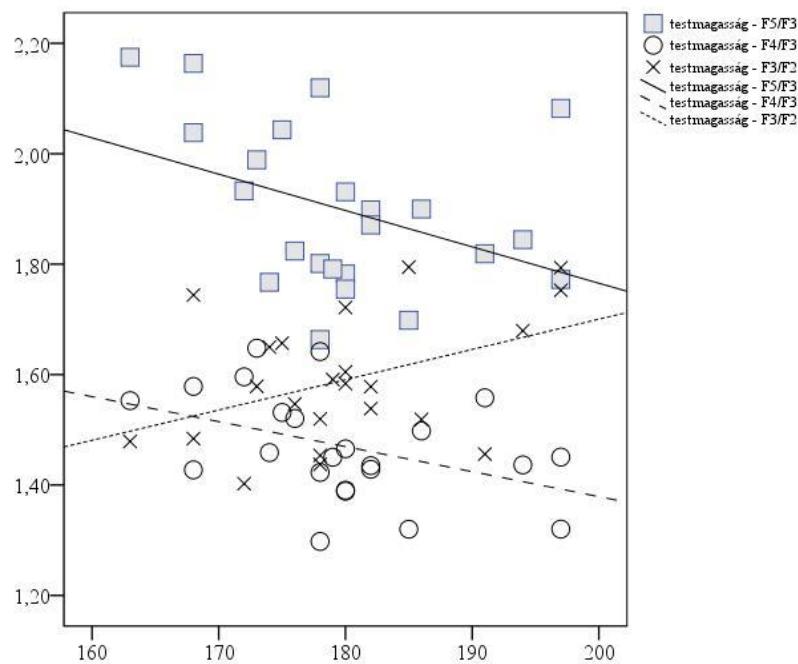
	<b>D<sub>1-2</sub></b>	<b>D<sub>1-3</sub></b>	<b>D<sub>1-4</sub></b>	<b>D<sub>1-5</sub></b>
<b>r</b>	-0,140	0,332	-0,197	-0,232

A következő lépésben az arányszámokkal végeztünk számításokat. Nem csak a Smith és munkatársai (2007) által javasolt F2/F1 és F3/F1 értékeit határoztuk meg, hanem az összes lehetséges arányszámot, amely a mért öt formánsérték között kiszámítható. Az 5. táblázatban csak három azt a adatot kö-

zöljük, ahol szignifikáns vagy tendenciaszerű (a  $p = 0,05$ -öt megközelítő szignifikanciaszinten) találtunk összefüggést. Három ilyen arányszám adódott, ezeket az 5. táblázat tartalmazza, grafikus ábrázolásuk pedig a 4. ábrán látható.

5. táblázat: A testmagasság és a formánsarányok korrelációs együtthatói

	F3/F2	F4/F3	F5/F3
<b>r</b>	0,418	-0,415	-0,397
<b>p</b>	0,047	0,049	0,061



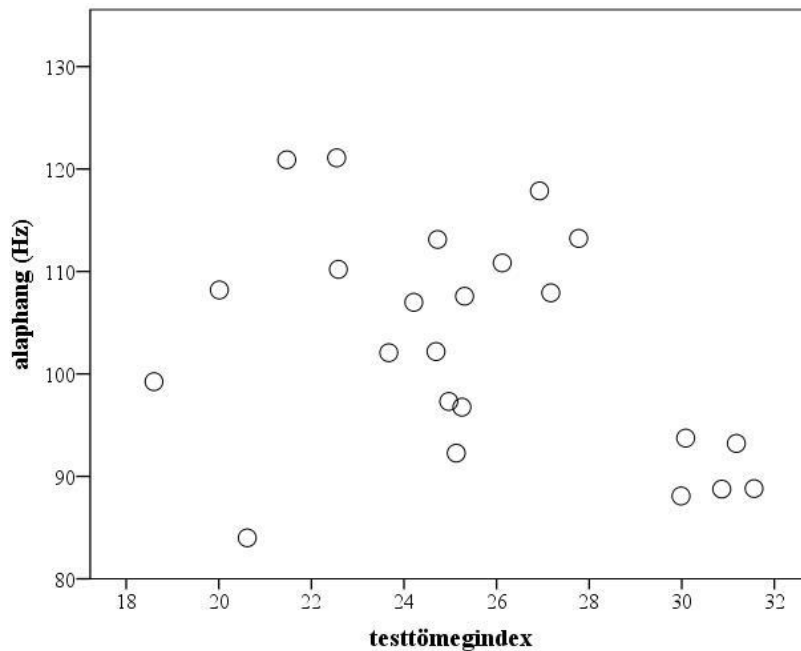
4. ábra

A formáns-arányszámok és a testmagasság összefüggései

Az ábráról leolvasható, de a korrelációs együttható előjele alapján is megállapítható, hogy az F3/F2 arány szignifikánsan nagyobb volt a magasabb beszélőknél. Másképp fogalmazva, a magasabb beszélőknél az F3 és az F2 egymástól távolabb, az alacsonyabb beszélőknél egymáshoz közelebb helyezkedett el. Az F4/F3 és az F5/F3 arányok ugyanakkor ellentétes irányú

kapcsolatot mutatnak a testmagassággal, a magasabb beszélőknél az F3-tól az F4, illetve tendenciaszerűen az F5 közelebb helyezkedett el. Ebből adódóan kisebb arányszám adódott.

Negyedik kutatási kérdésünk az akusztikai paraméterek és a testtömegindex közötti összefüggésre vonatkozott. Az 5. ábra mutatja az alaphang és a testtömegindex kapcsolatát.



5. ábra

A testtömegindex és az alaphang összefüggése

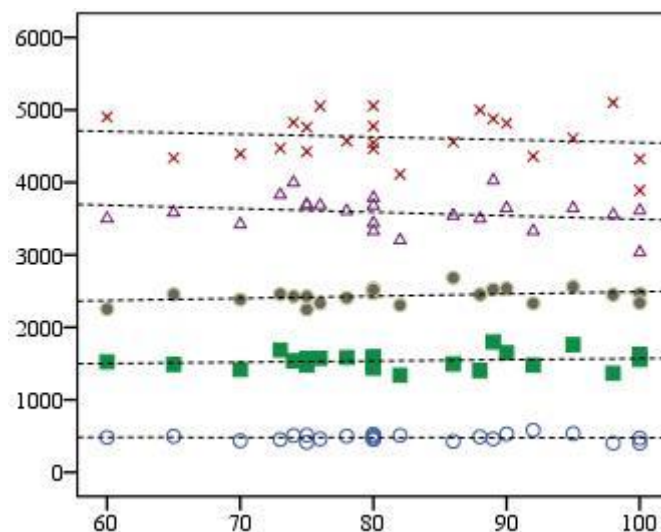
A korrelációs számítás nem igazolt szignifikáns korrelációt ( $r = -0,375$ ;  $p = 0,078$ ), de tendenciaszerűen a magasabb testtömegindexek mélyebb alaphanggal fordultak elő. Ha azonban a testtömegindex helyett a testsúllyal végezzük el a korrelációs számítást, akkor szignifikáns összefüggést kapunk ( $r = -0,584$ ;  $p < 0,01$ ), ami egyértelműen azt igazolja, hogy a magasabb testsúlyú személyek mélyebb alaphangon ejtették a kitöltött szüneteket.

A testtömegindex és a formánsstruktúra kapcsolatát is megvizsgáltuk. A 6. táblázat, illetve a 6. és 7. ábra adatai szerint nem mutatkozott szignifikáns összefüggés egyik esetben sem, és akkor sem, ha a fentiekhez hasonlóan, a számítást a testsúllyal végeztük el.

6. táblázat: A testtömegindex, a testsúly és a formánsok korrelációs együtthatói

		F1	F2	F3	F4	F5
testtömeg-index	$r$	0,100	0,311	0,121	-0,047	0,048
testsúly	$r$	-0,042	0,161	0,316	-0,229	-0,132

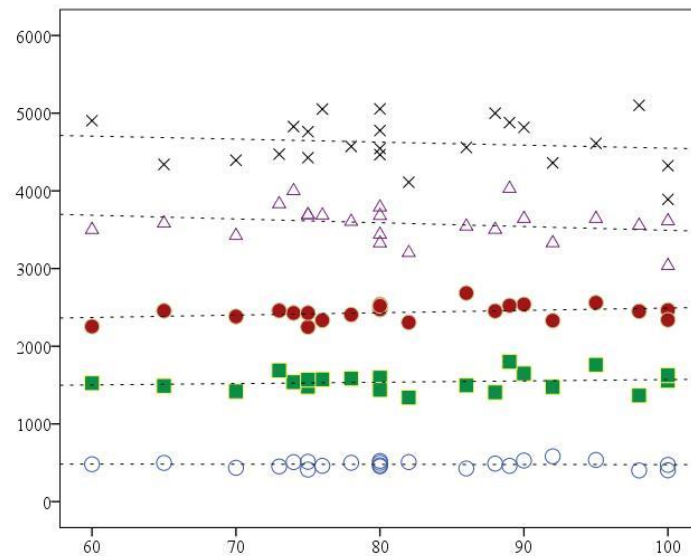
6. ábra



A testtömegindex és a formánsértékek összefüggései

Az ENTER módszerrel ismét megkíséreltünk regressziós modelleket felállítani. Ezek magyarázóereje a várakozásnak megfelelően nem volt szignifikáns (a testtömegindexre:  $F(5,17) = 0,961$ ; n.sz., a testsúlyra:  $F(5,17) = 1,321$ ; n.sz.).

Utolsó kutatási kérdésünk megválaszolásához először a mért formánsértékekből a (2) jelű képlet alapján megbecsültük a toldalékcső-hosszakat.



7. ábra

A testsúly és a formánsértékek összefüggései

Ezt követően korrelációs számítás segítségével megvizsgáltuk, hogy a becsült értékek összefüggnek-e a beszélők testmagasságával. A 7. táblázat tartalmazza a kapott korrelációs együtthatókat. Az adatok alapján megállapítható, hogy egyetlen esetben sem korrelált a becsült toldalékcsőhossz a formánsértékekkel.

7. táblázat: A formánsok alapján becsült toldalékcső-hosszak ( $L$ ) és a testmagasságok korrelációs együtthatói

	$L_{F1}$	$L_{F2}$	$L_{F3}$	$L_{F4}$	$L_{F5}$
$r$	0,255	0,253	-0,284	0,228	0,247

### Következtetések

Kutatásunk során azt vizsgáltuk, hogy az  $\bar{o}$ -féle kitöltött szüneteken mért alaphang, illetve első öt formáns összefüggést mutat-e a beszélő testmagasságával, illetve a testtömeg-indexével, testsúlyával.

Első és második kutatási kérdésünkre egy esetben sem kaptunk olyan választ, amely szignifikáns összefüggést igazolt volna, azaz, sem az alaphang, sem az első öt formáns értéke nem korrelált sem a testmagassággal, sem az életkorral. Így első hipotézisünk részben igazolódott – az alaphang nem kor-

relált a testmagassággal –, a második hipotézist viszont nem sikerült igazolni. Ennek ellenére a 2. ábrán olyan mintázat rajzolódik ki, amely azt sejteti, hogy az F3 kivételével a magasabb beszélőkhöz mélyebb formánsértékek tartozhatnak. Bár egyik kapcsolat sem szignifikáns, a korrelációs együtthatók elemzése bizonyos lehetőségeket előre vetít. Megállapítható ugyanis, hogy míg az F3-hoz tartozó korrelációs együttható gyakorlatilag 0-nak tekinthető, addig az F2-höz tartozó  $r$  értéke 0,328, ami abszolút értékben a legnagyobb a vizsgált öt formáns közül. Elképzelhető, hogy nagyobb és egységesebb mintán az F2 életkortól való függése igazolódik, azaz az idősebbeknél magasabb F2 mérhető. Ebben az esetben az  $\bar{o}$ -féle kitöltött szünetek ejtésével kapcsolatban hasonló megállapítások tehetők, mint amit Gósy és Bóna (2014) talált az [ɔ] hang esetében, tehát, az idősebbeknél a nyelv kissé előrébb helyezkedik el, emiatt adódik a magasabb F2. Ugyanakkor az F3 itt mért értékei inkább Gósy és Bóna (2014) kutatásában az [ɛ] magánhangzónál talált, közel azonos értékhez hasonlítanak.

A formánsdiszperzió értékei sem korreláltak a testmagassággal, ugyanakkor említést érdemel, hogy a  $D_{1-3}$  korrelációs együtthatójának értéke abszolút értékben jóval nagyobb a többinél ( $r = 0,332$ ). Ennek magyarázata lehet az, hogy az F3 és az F2 regressziós egyenesei kis mértékben széttartottak, így a magasabb beszélőkön nagyobb F3-F2 különbség tartozott (2. ábra). A formánsarányok vizsgálata azonban már szignifikáns kapcsolatot tárt fel. Úgy tűnik tehát, hogy – legalábbis az itt alkalmazott módszerrel gyűjtött adatok alapján – az F3 önmagában ugyan nem tükrözi a testmagasságot, de más formánsokkal képzett arányszámai szignifikáns kapcsolatban állnak a beszélő testmagasságával. Harmadik hipotézisünk tehát részben igazolódott, az egyik formánsarány-értékre szignifikáns összefüggés adódott.

Negyedik hipotézisünk is csak részben igazolódott. A nagyobb testsúlyú beszélők valóban mélyebb alaphangon beszéltek, a testtömegindex esetében azonban már nem adódott szignifikáns összefüggés, csak tendenciaszerű. A formánsadatok pedig egyik testalkati paraméterrel sem korreláltak.

Végül, ötödik hipotézisünket sem sikerült igazolni. A formánsértékekből kiszámított, megbecsült toldalékcso-hosszak egyik formáns esetében sem korreláltak a beszélő testmagasságával.

A feltárt eredmények a korábbi szakirodalmi adatokkal összevetve leginkább Hadman és munkatársai (2013) adatainak egy részével mutatnak hasonlóságot, akik az [a:] hang első négy formánsánál nem találtak korrelációt a beszélő testmagasságával. Igaz, kísérletükben [i:] hang második, harmadik és negyedik formánssa negatív előjellel, szignifikáns kapcsolatban állt a testmagassággal. A testtömegindex és a testsúly esetében Evans és munkatársai (2006) adataival megegyezően azt találtuk, hogy az alaphang nem korrelál a testmagassággal, a testsúly viszont igen. A formánsfrekvenciák alapján megbecsült toldalékcso-hosszak és a testmagasságok sem korreláltak. Ez az eredmény összhangban van Hatano és munkatársai (2012) eredményeivel,



akik öt magánhangzó közül négy esetben nem találtak korrelációt a formáns-szerkezet és a toldalékcso – MRI-felvételeken megmért – hossza között. Sőt, ellentmondva korábbi adatoknak, a toldalékcso hossza és a testmagasság között sem tudtak összefüggést igazolni.

Eredményeink alapján a legfontosabb kérdés annak tisztázása, hogy a kitöltött szünetek formánsadataiból miért nem lehetett egyértelműen következtetni a beszélők testmagasságára, vagy legalábbis miért nem adódott több helyen szignifikáns kapcsolat.

Ennek egyik oka az lehet, hogy a toldalékcsohossz és a testmagasság valóban nem, vagy csak nagyon gyengén korrelál (Hatano et al. 2012), így a várakozásokkal szemben az akusztikai adatokból nem, vagy csak igen nehezen, esetleg bizonyos feltételek mellett lehet a beszélő testméreteire visszakövetkeztetni.

A másik lehetséges okot a következőkben látjuk. Korábbi kutatások a leggyakrabban izolált hangzókat (Smith et al. 2007; Hatano et al. 2012; Hadman et al. 2013; Souza–Santos 2017) vagy jól meghatározott hangzókörnyezetben vagy egy szótagos szavakban ejtett magánhangzókat használták (Rendall et al. 2005; Pisanski–Rendall 2011; Pisanski et al. 2014a). Ezekben az esetekben a vizsgált magánhangzók nyelvi, pragmatikai funkciók, hangsúlyozás nélkül jelentek meg. Jelen kutatásunkban azonban spontán beszédből, különböző helyzetekben előforduló kitöltött szüneteket gyűjtöttük. Ismeretes, hogy a kitöltött szünetek megjelenése többféle okra vezethető vissza, és különböző funkciókban jelennek meg (Gósy et al. 2013, Horváth 2014:41, Laczkó 2013). Az egyes funkciókban megjelenő [ə]-szerű kitöltött szünetek formánsstruktúrái között kimutathatók különbségek, bár ezek a különbségek csak tendenciaszerűek (Horváth 2014: 44). Kutatásunkban nem különböztettük meg a kitöltött szüneteket ebből a szempontból, így elképzelhető, hogy az egyes beszélőknél különböző arányban fordultak elő a különböző nyelvi funkciókban megjelenő kitöltött szünetek. Ez pedig eredményezhette azt, hogy az egyes beszélőknél az adott beszédmintában gyakrabban előforduló funkciókban megjelenő kitöltött szünetek nagyobb súllyal szerepeltek az átlagértékekben, mint a ritkábban előforduló funkciók kitöltött szünetei.

További magyarázat lehet az is, hogy az egyes beszélők eltérő stratégiákat alkalmazhatnak az kitöltött szünetek realizációi során. Mivel spontán beszédet használtunk, a beszélők nem kaptak instrukciókat a kitöltött szünetek ejtéséhez. Így előfordulhat, hogy egyes beszélők egyéni sajátosságként erőteljesebb [ø]-s színezettel, mások hátrébb, [a:] -ba hajlóan képezték. Feltételeztük továbbá azt is, hogy az esetleges koartikulációs hatásokat ki tudjuk küszöbölni azzal, ha a formánsokat a hangzók közepén mérjük, illetve a mért 15-20 hangzó formánsértékeit átlagoljuk. Ezt azonban szintén nem ellenőriztük, például a különböző hangzókörnyezetekben előforduló kitöltött szünetek egymással való összehasonlításával, így az esetleges koartikulációs hatások sem zárhatók ki teljes bizonyossággal.

Az említett tényezők – a nyelvi funkció, az egyéni ejtési sajátosságok, illetve a koartikulációs hatások – torzító tényezőként szerepelhettek. Másképp fogalmazva, elképzelhető, hogy ha ezek nem jelentkeznének a realizáció során, akkor az így kapott, ideális, „tiszta” esetben az ő-féle kitöltött szünetek formánsszerkezete már jól tükrözné a beszélő testmagasságát. Valószínűsíthető az is, hogy nagyobb mintán az itt tendenciaszerűen jelentkező kapcsolatok is szignifikánsan kimutathatók lennének. Ennek ellenére célszerűbbnek tűnik a jövőben más magánhangzókra megismételni a méréseket, mivel azok jóval nagyobb számban állnak rendelkezésre, így spontán beszédből könnyebb megfelelő számú, hasonló tulajdonságú hangzót kigyűjteni a kísérletekhez.

### Irodalom

- Ahmadi, Farzaneh – McLoughlin, Ian 2012. Measuring resonances of the vocal tract using frequency sweeps at the lips. In *5th International Symposium on Communications Control and Signal Processing (ISCCSP)*.
- Babel, Molly – McGuire, Grant – King, Joseph 2014. Towards a More Nuanced View of Vocal Attractiveness. *PLoS ONE* 9/2. e88616. 1–10. doi:10.1371/journal.pone.0088616. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0088616> (A letöltés ideje: 2018. április 18.)
- Barreda, Santiago 2016. Investigating the use of formant frequencies in listener judgments of speaker size. *Journal of Phonetics* 55. 1–18.
- Boersma, Paul – Weenik, David 2014. *Praat: doing phonetics by computer [computer program]. Version 5.3.82.* <http://www.praat.org/> (A letöltés ideje: 2014. június 12.)
- Bóna Judit – Vakula Tímea 2013. Temporális sajátosságok 90 évesek spontán beszédében. *Alkalmazott Nyelvstudomány* XIII/1-2. 59–73.
- Charlton, Benjamin D. – Taylor, Anna M. – Reby David. 2013. Are men better than women at acoustic size judgements? *Biology Letters* 9. <http://dx.doi.org/10.1098/rsbl.2013.0270> (A letöltés ideje: 2018. április 18.)
- van Dommelen, Wim A. – Moxness, Bente H. 1995. Acoustic parameters in speaker height and weight identification: Sex-specific behavior. *Language and Speech* 38/3. 267–287.
- Dusan, Sorin 2005. Estimation of speaker's height and vocal tract length from speech signal. *Interspeech 2005*. 1989–1992.
- Evans, Sarah – Neave, Nick – Wakelin, Delia 2006. Relationships between vocal characteristics and body size and shape in human males: An evolutionary explanation for a deep male voice. *Biological Psychology* 72. 160–163.
- Fant, Gunnar 1966. A note on vocal tract size factors and non-uniform F-pattern scalings. *Quarterly Progress and Status Report* 7/4. 22–30.
- Fant, Gunnar 1967. Sound, features and perception. *Quarterly Progress and Status Report* 8/2-3. 1–14.
- Feinberg, David R. – Jones, Benedict C. – Little, Anthony C. – Burt, Michael D. – Perrett, David I. 2005. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour* 69. 561–568.

- Fitch, Tecumseh, W. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of Acoustical Society of America* 102/2. 1213–1222.
- Fitch, Tecumseh, W. – Giedd, Jay 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*. 106/3. 1511–1522.
- Fitch, Tecumseh, W. – Reby, David 2001. The descended larynx is not uniquely human. *Proceedings of the Royal Society, Biological Sciences* 268. 1669–1675.
- Gianni, Antonella – Pettorino, Massimo – Cinque, Umberto 1989. Speakers' identification by voice. *Eurospeech* 89. 1283–1286.
- Gocsál Ákos 2017. Az artikulációs tempó és az átlagos alaphang szerepe a beszélő életkorának megbecslésében. *Beszéd kutatás 2017*. 151–167.
- González, Julio 2003. Estimation of speakers' weight and height from speech: A re-analysis of data from multiple studies by Lass and colleagues. *Perceptual and Motor Skills* 96. 297–304.
- González, Julio 2004. Formant frequencies and body size of speaker: a weak relationship in adult humans. *Journal of Phonetics* 32. 277–287.
- González, Julio 2007. Correlations between speakers' body size and acoustic parameters of voice. *Perceptual and Motor Skills* 105. 215–220.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125. 137–148.
- Gósy Mária 2006. A semleges magánhangzó nyelvi funkciói. *Beszéd kutatás 2006*. 8–22.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácz Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Gósy Mária – Bóna Judit – Beke András – Horváth Viktória 2013. A kitöltött szünetek fonetikai sajátosságai az életkor függvényében. *Beszéd kutatás 2013*. 121–143.
- Gósy Mária – Bóna Judit 2014. Magánhangzók ejtése fiatalok és idősek spontán beszédében. *Magyar Nyelv* 110/2. 129–143.
- Gyarmathy Dorottya 2015. Diszharmóniás jelenségek, megakadások a beszédben. In Gósy Mária (ed.) *Diszharmóniás jelenségek a beszédben*. MTA Nyelvtudományi Intézet, Budapest. 9–47.
- Hadman, Abdul-Latif H. – Al Barazi, Randa – Khneizer, Gebran – Turfe, Zaahir – Ashkar, Jidah – Tabri, Dollen 2013. Formant frequency in relation to body mass composition. *Journal of Voice* 27/5. 567–571.
- Hatano, Hiroaki – Kitamura, Tatsuya – Takemoto, Hironori – Mokhtari, Parham – Honda, Kiyoshi – Masaki, Shinobu 2012. Correlation between vocal tract length, body height, formant frequencies, and pitch frequency for the five Japanese vowels uttered by fifteen male speakers. *Interspeech 2012*. 402–405.
- Hayward, Katrina 2000. *Experimental phonetics*. Longman Linguistics Library. Pearson Education Limited, Harlow.
- Horváth Viktória 2014. *Hezitációs jelenségek a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- Johnson, Keith 2003. *Acoustic & Articulatory Phonetics*. Blackwell Publishing, Oxford.

- Krauss, Robert M. – Freyberg, Robert – Morsella, Ezequiel 2002. Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology* 38. 618–625.
- Künzel, Hermann J. 1989. How well does average fundamental frequency correlate with speaker height and weight? *Phonetica* 46. 117–125.
- Laczko Mária 2013. A kitöltött szünetek formái és funkciója tizenévesek spontán beszédében. *Magyar Nyelvőr* 137/2. 192–208.
- Lindblom, Björn E. F. – Sundberg, Johan 1972. Approaches to articulatory modeling. *GALF* 1972. 3–45. [http://www.afcp-parole.org/doc/Archives\\_JEP/1972\\_IIIe\\_JEP\\_Lannion/1972%20-%20IIIe%20JEP%20-%20Lannion.pdf](http://www.afcp-parole.org/doc/Archives_JEP/1972_IIIe_JEP_Lannion/1972%20-%20IIIe%20JEP%20-%20Lannion.pdf) (A letöltés ideje: 2018. április 18.)
- Markó Alexandra 2013. *Az irreguláris zöngé funkciói a magyar nyelvben*. ELTE Eötvös Kiadó, Budapest.
- Pisanski, Katarzyna – Rendall, Drew 2011. The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America* 129/4. 2201–2212.
- Pisanski, Katarzyna – Fraccaro, Paul J. – Tigue, Cara C. – O'Connor, Jillian J. M. – Feinberg, David R. 2014a. Return to Oz: Voice pitch facilitates assessments of men's body size. *Journal of Experimental Psychology: Human Perception and Performance* 40/4. 1316–1331.
- Pisanski, Katarzyna – Fraccaro, Paul J. – Tigue, Cara C. – O'Connor, Jillian J. M. – Röder, Susane – Andrews, Paul W. – Fink, Bernhard – DeBruine, Lisa – Jones, Benedict C. – Feinberg, David R. 2014b. Vocal indicators of body size in men and women: a meta-analysis. *Animal Behaviour* 95. 89–99.
- Puts, David Andrew – Hodges, Carolyn R. – Cárdenas, Rodrigo A. – Gaulin, Steven J. C. 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior* 28. 340–344.
- Rendall, Drew – Kollias, Sophie – Ney, Christina – Lloyd, Peter 2005. Pitch ( $F_0$ ) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America* 117/2. 944–955.
- Smith, David R. R. – Patterson, Roy D. 2005. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America* 118/5. 3177–3186.
- Smith, David R. R. – Walters, Thomas C. – Patterson, Roy D. 2007. Discrimination of speaker sex and size when glottal-pulse rate and vocal-tract length are controlled. *The Journal of the Acoustical Society of America* 122/6. 3628–3639.
- Souza de, Lourdes Bernadete Rocha – Santos dos, Marquiony Marques 2017. Body mass index and acoustic voice parameters: is there a relationship? *Brazilian Journal of Otorhinolaryngology* <http://dx.doi.org/10.1016/j.bjorl.2017.04.003> (A letöltés ideje: 2018. április 18.)

**An analysis of correlations between some body size parameters and the formant structure of filled pauses in male speakers**

The main purpose of this paper is to examine if correlations exist between body height, body mass index, and the formant values of filled pauses in male speakers. Based on acoustic theories of speech, several mathematical formulas exist, which explain the formant frequencies of vowels. It has been suggested by several authors that formants may be good predictors of body size, especially that of body height. The experimental results obtained so far are, however, not unambiguous. For the present experiment, schwa-like filled pauses were extracted from spontaneous speech recordings of 23 male speakers and fundamental frequencies and the values of the first five formants were measured. None of these parameters correlated with body height, although a nonsignificant tendency of taller speakers having lower formant frequencies, except for F3, was observed. Body mass index did not correlate with formant frequencies and fundamental frequency values, however, heavier speakers used significantly lower f0. Vocal tract lengths were also estimated, based on the formant frequencies, but they did not correlate with actual speaker heights.

**MONDATHANGSÚLYOS ÉS HANGSÚLYTALAN  
HELYZETŰ MAGÁNHANGZÓK  
NÉHÁNY ARTIKULÁCIÓS ÉS AKUSZTIKAI  
JELLEMZŐJE A MAGYARBAN**

**Markó Alexandra<sup>1,4</sup> – Bartók Márton<sup>1,4</sup> –  
Grácsi Tekla Etelka<sup>2,4</sup> – Deme Andrea<sup>1,4</sup> –  
Csapó Tamás Gábor<sup>3,4</sup>**

<sup>1</sup> ELTE Fonetikai Tanszék, <sup>2</sup> MTA Nyelvtudományi Intézet, <sup>3</sup> BME TMIT,  
<sup>4</sup> MTA-ELTE Lendület Lingvális Artikuláció Kutatócsoport

**Bevezetés**

Prominencia (kiemelés) a nyelvi viselkedés különféle aspektusaiban megfigyelhető, és változatos eszköztárral valósulhat meg. Attól függően, hogy milyen nyelvelmélet alapján közelítjük meg, illetve hogy a nyelv mely rétegében vizsgáljuk, számos definíció és vizsgálati módszertan állhat rendelkezésünkre. A hangtani (fonetikai-fonológiai) leírások a prominenciát általában a hangsúly ekvivalenseként, illetve a hangsúlyt a prominencia kifejezőjeként tartják számon (pl. Kálmán–Nádasdy 1994/2016; Gósy 2004; Fletcher 2010; Mády 2012). Ugyanakkor (a magyarban a hangsúlytól nem függetlenül) prominenciát fejezhet ki a szórend is, ahogyan a szintaxis, a szemantika és a pragmatika felől nézve is megkülönböztethetünk prominensebb és kevésbé prominens elemeket (vö. pl. Imrényi 2007).

A jelen tanulmány – hangtani vonatkozású lévén – a prominenciát és a hangsúlyhelyzetet ekvivalensként kezeli a továbbiakban, és a hangsúlyt perceptuális kategóriaként ragadja meg, amelynek az artikulációs, illetve akusztikai eszközökkel történő megvalósítását vizsgálja. A fonetikában a hangsúly akusztikai korrelátumainak tekintik általánosságban az időtartam, az intenzitás, az alaphékvencia és a spektrális szerkezet valamiféle többletét vagy eltérését a hangsúlyos szótagban – a környező hangsúlytalanokhoz képest –, illetőleg a hangsúly ugyancsak összefüggést mutat a fonáció típusával és a vokális erőfeszítéssel (Marasek 1996). Mindezeknek az akusztikai és artikulációs jellemzőknek az elemzése ugyanakkor rendre ellentmondásos eredményekre vezet (lásd az alábbi szakirodalmi áttekintésben).

Markó Alexandra – Bartók Márton – Grácsi Tekla Etelka – Deme Andrea – Csapó Tamás Gábor 2018. Mondathangsúlyos és hangsúlytalan helyzetű magánhangzók néhány artikulációs és akusztikai jellemzője a magyarban. *Beszédkutatás* 2018. 85–109.

A hangsúly hatókörét tekintve szokás megkülönböztetni szóhangsúlyt és mondathangsúlyt, ugyanakkor a szószintű és a megnyilatkozásszintű hangsúly nem függetleníthető egymástól. Tekintettel arra, hogy a beszéd megnyilatkozásokból áll, a szavakon belüli hangsúlyviszonyokról a megnyilatkozásbeli helyzetüktől függetlenül nem is igen tehetők megállapítások. Általánosságban a magyar nyelvben a szóhangsúly helyzete kötött, az első szótagon realizálódik, illetve az első szótagi szóhangsúly részletező jelentésszerkezetű (vö. Tolcsvai Nagy 2017: 208; más megközelítésben alapszófajú, vö. Keszler szerk. 2000) szavak esetében jelentkezik, míg a sematikus jelentésszerkezetű szavak (pl. névelők, névutók, kötőszók; vö. Tolcsvai Nagy uo.) klitikumként realizálódnak, azaz nem jelenik meg szóhangsúly egyik szótagjukon sem. Ha azonban a megnyilatkozásban szemantikai és/vagy pragmatikai okok indokolják (vö. még ún. ingatag hangsúlyhelyzetek, Kálmán–Nádasdy 1994/2016), a részletes jelentésszerkezetű szavak első szótagjáról el is mozdulhat a hangsúly, és akár teljesen hangsúlytalaná is válhatnak (vö. irtóhangsúly, Kálmán–Nádasdy 1994/2016 és lásd alább). Ugyancsak hangsúlyossá válhatnak ilyen okok miatt a sematikus jelentésszerkezetű szavak (például kontraszt kifejezésében).

A megnyilatkozás szintjén a hangsúly többféleképpen realizálódhat a magyar beszédben. Tág fókusz esetén minden prozódiai (részletes jelentésszerkezetű) szóhangsúlyt kap, a szűk fókuszú megnyilatkozásban azonban a fókuszpozícióban álló szó a legkiemeltebb, míg az összes többi hangsúlytalanodik (irtóhangsúly). Ez utóbbi esetben a fókusz meghatározott helyen szerepel a magyar mondatban (az ige előtt). A szintaxis és a hangsúlyelosztás szoros kapcsolata miatt felmerült a kérdés, hogy valóban alkalmazunk-e prozódiai eszközöket a fókusz jelölésére a magyar nyelvben. Ezek hiányát több tanulmány alátámasztotta mind laboratóriumi, mind spontán beszédben (pl. Mády–Kleber 2010; Markó 2012), míg mások igazolták meglétüket például az alapfrekvencia ( $f_0$ ) maximumában és terjedelmében (Genzel et al. 2015 és hivatkozásai).

A magyar mellékhangsúly létezésének kérdése több szempontból is vitákat váltott ki (vö. pl. Varga 1985, 2000, 2012; Kálmán–Nádasdy 1994/2016; Blaho–Szeredi 2011 és hivatkozásaik). Varga (1985) a magyarban három hangsúlyfokozatot különít el: főhangsúlyt, mellékhangsúlyt és zérushangsúlyt. Kálmán és Nádasdy (1994/2016) szerint a mellékhangsúly csak fonetikai, de nem fonológiai entitás. Mellékhangsúlyt Varga (2000: 93) négyféle pozícióban feltételez: összetett szó nem első tagjának első szótagján; deakcentuált szavak első szótagján; több szótagú funkciószavak első szótagján; valamint bizonyos szó belseji (rendszerint páratlan) szótagokon ritmikai okokból. Blaho és Szeredi (2011) a ritmikai mellékhangsúly létét akusztikai vizsgálataik alapján elvetették, a többi, Varga által felsorolt esetet nem vizsgálták. A vita további részletei a jelen tanulmány szempontjából irrelevánsak, itt csak azt jegyezzük meg, hogy az általunk vizsgált anyagban a

hangsúlytalanként minősített szótagok a fenti kritériumrendszer alapján is hangsúlytalanok, (feltételezett) mellékhangsúllyal sem rendelkeznek.

A magánhangzók minősége a tudományos közvélekedés szerint a magyar beszédben nem változik a hangsúly (sem a mondathangsúlyt nem viselő, sem a mondathangsúllyal egybeeső szóhangsúly) jelenlétének függvényében (vö. pl. Kálmán–Nádasdy 1994/2016; Gósy 1997 és hivatkozásai). Ugyanakkor ez az állítás tapasztalati (percepciós) alapú, hiszen erre vonatkozóan (a jelen tanulmány szerzőinek tudomása szerint) nem állnak rendelkezésre szisztematikus összevetésből származó, azaz például a hangkörnyezet, illetve a szótag-szerkezet azonossága mellett elvégzett mérésekből származó adatok. Az ugyanakkor elképzelhető, hogy miközben fonológiai relevanciájú minőségi és mennyiségi eltérés nincs ugyanazon magánhangzó hangsúlyos és hangsúlytalan szótagbeli realizációi között, a hangsúlyosság hordozója mégis valamiféle fonetikai (akusztikai és/vagy artikulációs mérésekkel detektálható) különbség.

Fónagy nevezetes, máig sokat idézet tanulmánya (1958) a hangsúlyosságot az időtartam, a(z érzeti) hangosság és az alapfrekvencia relatív többletével jellemzi, amelyekkel kapcsolatosan többször megjegyzi, hogy ezek közül bármelyik hiányozhat, és a kapcsolatuk a hangsúllyal csak tendenciaszerű. Akusztikai mérések mellett Fónagy fiziológiai vizsgálatokat is végzett gége-mikrofonnal, pneumográffal, elektromiográffal, amelyek eredményei alapján arra a következtetésre jutott, hogy a hangsúly elsősorban a belső bordaközi izmok, másodsorban a gégeizmok tevékenységével hozható összefüggésbe, miközben a nagyobb izomaktivitás egyáltalán nem feltétlenül vezet mérhető akusztikai (pl. intenzitásbeli) többlethez. Fónagy a kísérletek hozadékát így összegzi: „A hangsúlyélmény [...] nem a hallószervben, hanem a központi idegrendszerben alakul ki a hangjelenség alapos, de gyors és öntudatlan elemzése nyomán. [...] mindenki saját tapasztalatából tudja, hogyan tükröződik a beszédben az erő kifejtés. A hallottak alapján követni tudjuk a beszélő fiziológiai tevékenységét” (23–24). Ez a konklúzió egyrészt összecseng a tükrőneuronokkal kapcsolatos vizsgálatok eredményeiből levont azon következtetésekkel (vö. pl. Thèoret–Pascual-Leone 2002), amelyek szerint azért értjük meg mások, így a beszédpartnerek cselekvéseit, azaz például a nyelvi viselkedésüket és ezen belül a prominencia kifejezését is, mert e cselekvések megfigyelése ugyanazokat az idegsejteket aktiválja, mint ezeknek a cselekvéseknek a kivitelezése. Másrészt Fónagy megállapítása ugyancsak harmonizál a beszédészlelés motoros elméletében foglalt azon állítással, amely szerint a beszéd észlelése az akusztikai jel transzformációján keresztül az artikulációs mozdulatsor reprezentációjának előhívásával megy végbe (Lieberman–Mattingly 1985).

A hangsúlyosság elemzése a következő évtizedekben nem állt a magyar beszédre vonatkozó kísérleti fonetikai vizsgálatok előterében, de mint szempontot rendre figyelembe vették a szerzők a magánhangzók vizsgálatában.



Magdics (1965) monográfiája felolvasott mondatok magánhangzóinak átlagos formánsfrekvencia-adatait sorjázza. Az adatokat a szerző különválasztotta a hangsúlyos és a hangsúlytalan magánhangzók tekintetében; az ő megfogalmazásában: „Hangsúlyos és hangsúlytalan magánhangzón [...] hangsúlyos és hangsúlytalan helyzetbeli magánhangzót értek” (6), azt azonban nem határozza meg, hogy ezek mit jelentenek pontosan. Azt gondolhatjuk, hogy a szavak szintjén első és nem első szótagi helyzetekre utal így a szerző, és feltehető, hogy a grammatikai funkciójú lexémák (vö. sematikus jelentésszerkezetű szavak) esetében hangsúlytalan szótagokról van szó az első szótag esetében is. Az adatok alapján Magdics azt állapítja meg, hogy a hangsúlytalan helyzetben „a magánhangzók formánsai lefelé tolódnak” (21), ami ebben a formában azt jelentené, hogy a magánhangzótér egységesen a zártabb és a hátrébb képzettség irányába mozdul, azonban a részletes elemzésből az is kiderül, hogy a tendencia nem egységes. Általánosságban valamivel periferikusabb ejtésre utalnak a hangsúlytalan helyzetben mért formánsértékek, habár a számszerűsített eltérések statisztikai elemzése nem történt meg. Mivel a hangkörnyezetek és a mondatbeli pozíciók is igen változatosak (például nincs kontrollálva az esetleges fókuszhangsúly), és a beszélők is váltakoznak (a felvételek 5 bemondóval készültek, de a ritka vagy az anyagból hiányzó hangzók esetében Magdics a saját beszédét elemezte, amelyet más körülmények között rögzített, mint a többi bemondó esetében), így szisztematikus összevetésre ezek az adatok nem adnak módot. Ugyanezen az anyagon mért időtartamadatok is publikált Magdics (1966), ezek szerint az átlagértékeket tekintve a hangsúlyosnak besorolt magánhangzók valamivel hosszabbak voltak, mint a hangsúlytalanok (vö. 2. táblázat, 130–131). Ez az eredmény elöntmond például annak a Lindblom (1963) modelljében megfogalmazott állításnak, mely szerint a realizáció rövidebb időtartama azt idézi elő, hogy a beszédsszerveknek „nincs idejük” elérni a célkonfigurációt, ezért célalulmúlás történik, szemben a hosszabb időtartamban realizálódó beszédhangokkal, ahol a célkonfiguráció megvalósul. Ennek értelmében az artikulációs szervek mozdulataiból következően a rövidebb beszédhangok esetén a formánsértékek centralizáltabb ejtésre utalnának, míg a hosszabb időtartammal periferikusabb ejtésre utaló formánsértékek járnának együtt. Magdics (1965, 1966) adatai azonban éppen ezzel ellenkező tendenciára engednek következtetni.

Kassai (1979) a magánhangzók időtartamát a hangsúlyos és a hangsúlytalan helyzet, a mássalhangzó-környezet és a szótagszerkezet függvényében elemezte szólistaként felolvasott, illetve mondatokba ágyazott minimális párokban, három beszélő ejtésében. Eredményei (1. és 2. táblázat, 38–39) arra utalnak, hogy az időtartambeli eltérések iránya és mértéke magánhangzó-minőségenként változó. A szerző véleménye szerint a hangsúlytalanság az értékek nagyobb szóródását eredményezi, míg a hangsúlyos szótag „igyekszik megőrizni az érvényben lévő fonetika-fonológiai állapotot” (42). Hasonló eredményeket kapott a spontán beszéd vizsgálatában is (4. táblázat, 64).

Gósy (1997) egy beszélő spontán beszédbeli magánhangzóit elemezte akusztikai szempontból, illetve percepció vizsgálatot is végzett, amelynek a kérdése az volt, hogy a különféle magánhangzó-minőségeket minek azonosítják az adatközlők: az eredeti magánhangzó-minőségnek, vagy pedig az /ø/-nek (a módszertan további részleteire vonatkozóan lásd az eredeti forrást). Az utóbbi esetben Gósy azt feltételezte, hogy a beszédben [ə] hangzott el. A szerző a hangsúlyos (azaz első szótagi) és hangsúlytalan (azaz nem első szótagi – a részletező és a sematikus jelentésszerkezetű szavak esetleges megkülönböztetésére; illetve az esetleges megnyilatkozásszintű hangsúly kezelésére nem tér ki) realizációk összevetése során változatos eredményeket kapott. Míg általában a hangsúlytalan szótagban ejtett magánhangzót azonosították sváként az adatközlők, az /o/ ettől eltérő tendenciát mutatott. Ez egyebek között a magánhangzó-minőségek közötti eltérésre is visszavezethető lehet. A magánhangzó-időtartamok ugyancsak különféleképpen tértek el az egyes magánhangzó-minőségek esetében, egyes esetekben a hangsúlyos, más esetekben a hangsúlytalan realizációk voltak átlagosan hosszabbak.

Kovács (2002) egy női adatközlő által háromszor felolvasott mondatokban több szempont mentén vizsgálta a magánhangzó-időtartamokat. Szisztematikusan párokba rendezett szavakat tartalmazó anyagában a hangsúlytalan szótagokban a szerző átlagosan 8%-nyi időtartamtöbbletet mért (70), mely eredmény részben ellentmondásban van azokkal a korábbiakkal, melyek a hangsúlyos szótag időtartamtöbbletére utalnak (Magdics 1966).

Mády és munkatársai (2008) tanulmányának szintén nem a hangsúlyosság hatásának vizsgálata volt az elsődleges célja, a kutatás mégis hozott e tekintetben is releváns eredményeket. A szerzők három rövid és hosszú magánhangzó-pár tagjainak megvalósulását elemezték azonos mássalhangzó-környezetben, 10 beszélő felolvasásában. Ebben a kutatásban ugyancsak magánhangzónként eltérő tendenciákat találtak, mind a magánhangzó-minőség, mind az időtartam tekintetében a hangsúlyos és a hangsúlytalan szótagbeli realizációk között.

Gósy és Beke (2010) 5 férfi és 5 nő spontán beszédében elemezte a magánhangzók időtartamát, megkülönböztetve az első és az utolsó, valamint a közbülső szótagi magánhangzókat (ennél közelebbi összefüggést a hangsúlylyal, azaz hogy a szó- vagy a megnyilatkozásszintű hangsúlyjellemzőkre reflektálnak-e az adatok, nem tartalmaz a tanulmány). Eredményeik szerint a vizsgált anyagban a közbülső szótagbeli magánhangzókétől mind az első, mind az utolsó szótagi magánhangzók időtartama pozitív irányban (szignifikánsan) eltért, ami részben összefügghet azzal is, hogy az első szótagi magánhangzók jobbára – részletező jelentésszerkezetű szavak esetén – (szó)hangsúlyos helyzetűek, és a szóhangsúllyal időtartamtöbblet jár(hat)ot együtt.

Szalontai és munkatársai (2016), illetőleg Mády és kollégái (2017) végezték a legújabb akusztikai elemzéseket a magyar hangsúly témakörében. Kuta-

tásuk célja a szó- és a mondathangsúly akusztikai vetületének elkülönítése volt a magyarban és a németben, kontrasztív vizsgálat keretében. Az elemzett (azonos fonológiai minőségű) magánhangzók szóhangsúlyos, szó szinten hangsúlytalan, mondathangsúlyos és mondatszinten hangsúlytalan pozícióban jelentek meg, mindig ugyanabban a szótagban. A magyarra vonatkozóan 1. +szóhangsúly, +mondathangsúly; 2. +szóhangsúly, –mondathangsúly és 3. –szóhangsúly, +mondathangsúly (ebben a kondícióban a vizsgált szótag nem abszolút szókezdő pozícióban állt) helyzeteket vizsgáltak és hasonlítottak össze páronként (1. vs. 2. és 1. vs. 3.). Az időtartamot, az energia négyzetes összegének maximumát, a spektrális egyensúlyt (a felső frekvenciatartomány energiája a teljes spektrumon mért hangnyomás értékével csökkentve) és az  $f_0$ -maximumot elemezték 12 beszélő kétszeri ejtésében az /o/ magánhangzóban. A két tanulmány (Szalontai et al. 2016 és Mády et al. 2017) vizsgálati módszertana némiképp eltér, az utóbbiban prozódiai stilizációval nyert adatokat elemeztek. Eredményeik szerint az időtartam és a normalizált  $f_0$ -maximum mind a szóhangsúly, mind a mondathangsúly jelölésében, míg a normalizált energiamaximum és az  $f_0$ -maximum csak a mondathangsúly realizálásában játszott szerepet.

Ugyan – mint láttuk – több kutatás foglalkozott a magyar beszédben a hangsúly fonetikai megvalósulásával, ezek legtöbbször az akusztikai szerkezetet vizsgálta. Az ötvenes évek óta nem keletkezett újabb, a magyar hangsúly jelenségét artikulációs alapon elemző tanulmány – a szerzők tudomása szerint. Ugyanakkor a nemzetközi szakirodalomban vannak olyan frissebb kísérletek, amelyek a prominencia és az artikuláció összefüggéseit vizsgálják. Az alábbiakban ezek közül a hangszalagműködésre vonatkozókból ismertetünk néhányat, mivel a nagyobb vokális erőfeszítés mint a hangsúly artikulációs korrelátuma gyakran képezi vizsgálatok tárgyát.

A hangszalagok rezgésének egy ciklusa háromfázisú folyamatnak tekinthető (vö. Marasek 1996): 1. nyitódó fázis, melynek során a hangszalagok távolodnak egymástól, így fokozatosan növekszik a hangrésen kiáramló levegő mennyisége; 2. záródó fázis, melynek során a hangszalagok közelítenek, ezáltal csökken a kiáramló levegőmennyiség; 3. zárt fázis, mely közben a hangszalagok között nincs levegőkiáramlás. Bizonyos fonációs típusok (pl. leheletes zöngé, suttogás), kismértékű vokális erőfeszítés, illetve bizonyos beszélők esetén nem különíthető el zárt fázis. A nagyobb vokális erőfeszítés a hangszalagok gyorsabb záródásával és hosszabb zárt fázissal jár együtt (Mooshammer 2010). A hangsúly és a nagyobb vokális erőfeszítés között a nemzetközi szakirodalom alapján összefüggést tételezhetünk, a hangsúly jelölése/megjelenése pedig ilyenformán várhatóan a zöngékepzés folyamatában és a zöngé minőségében is tetten érhető. Az eddigi kutatások azonban olyan nyelveket vizsgáltak elsősorban (például a németet, az angolt és a hollandot), amelyek a magyartól merőben eltérnek például a szintaktikai szerkezet és a lexikai hangsúly mintázataiban.

A hangszalagok működésének vizsgálata elektrolottográfia (rövidítése EGG) vagy más néven laringográfia történik. Ennek alkalmazásakor egy szalag segítségével két elektródát illesztenek a nyak két oldalára, a pajzsporc régiójában. Az elektródák minimális erősségű áramot bocsátanak ki, és az eszköz a gége és a környéki szövetek ellenállását méri. Az elektromos jel hullámformájából (lásd az 1. ábra EGG-vel jelölt paneleit alább, a módszertani fejezetben) megbecsülhető a hangszalagok érintkezésének és szétválásának pillanata. A jel legmagasabb értéke utal a hangszalagok maximális érintkezésére, az alacsony értékek pedig a nyitott állapotot mutatják. Az EGG-jel alapján többféle elemzés végezhető. Ezek közé tartozik a nyitott vagy a zárt állapot időtartamának arányítása a teljes fonációs periódusidőhöz. A nyitott állapottal történő számítást angol elnevezése (*open quotient*) alapján OQ-ként szokás rövidíteni. (A továbbiakban ezt alkalmazzuk.) A nagyobb időarányban nyitott hangszalagok és az ennek következtében mérhető magasabb OQ-értékek a modálishoz képest levegősebb zöngéképzésre engednek következtetni (vö. Kankare et al. 2012; Childers–Lee 1990), míg az alacsonyabb OQ-értékek nagyobb időtartamú hangszalag-érintkezéssel járó fonációs típusokra (például a préselt zöngére – vö. Kankare et al. 2012 – vagy az irreguláris zöngéképzés bizonyos típusaira – vö. Childers–Lee 1990), illetőleg a nagyobb vokális erőfeszítésre (vö. Mooshammer 2010) jellemzők.

Egy, a német nyelvre végzett tanulmány (Marasek 1996) 5 feszes és 5 laza magánhangzóban azt találta, hogy az OQ eltért a magánhangzó feszessége és hangsúlyossága mentén. A szerző azt találta, hogy a hangsúlyos magánhangzók OQ-értékei 3%-kal magasabbak voltak a hangsúlytalanokénál, ami egyébként eltér sok más eredménytől (ez utóbbiak összefoglalását lásd Mooshammer 2010). Marasek (1996) a szóhangsúlyt vizsgálta, de mivel szólisták felolvasását elemezte, bizonyos esetekben a szó- és a megnyilatkozásszintű hangsúly egybeeshetett, míg máshol eltérő szótagokon realizálódhatott – részben ez állhatott a más kutatások eredményeitől való eltérések hátterében. Ezen túlmenően az adatközlők nemét is figyelembe véve Marasek (1996) azt találta, hogy a hangsúly és a beszélő nemének hatása között interakció figyelhető meg: míg a férfiaknál nem volt eltérés az egyes vizsgált helyzetek között, a nők esetében a hangsúlyos szótagbeli magánhangzókra kapott OQ-érték átlagosan 7%-kal magasabb volt, mint a hangsúlytalan helyzetben mért adatok. Ez arra utalhat, hogy a vizsgálat női adatközlői a hangsúlyos helyzetű magánhangzókat levegősebb zöngéképzéssel ejtették, mint a hangsúlytalanokat.

Mooshammer (2010) vizsgálatában az /e/ megvalósulásait elemezte a *le* szótagban a német nyelvben (7 férfi ejtésében). A célmagánhangzók négy prominenciafokban szerepeltek a vizsgálati anyagban úgy, hogy az azokat tartalmazó szótagok a szóhangsúly megléte és hiánya, illetve a mondathangsúly megléte és hiánya mentén variálódtak rendszerszerűen. Emellett a szerző a mondat- és szóhangsúlyos helyzetben a „hangos” és a „halk” ejtést is fel-

vette változóként (ehhez mindösszesen annyi utasítást adott a beszélőknek, hogy a célszavakat ejtsék halkán vagy hangosan). Több akusztikai és artikulációs paraméter közül az OQ bizonyult általában a hangsúly legkövetkezetesebb velejárójának: Mooshammer átlagosan 2,5%-kal alacsonyabb OQ-értékeket mért a hangsúlyos szótagokban, mint a hangsúlytalanokban – ez nagyobb vokális erőfeszítésre utal a hangsúlyos ejtés esetében.

A két tanulmányt összevetve azt láthatjuk, hogy azok eltérő eredményeket találtak az OQ és a hangsúly összefüggésében. Mooshammer (2010) ezt egyebek mellett azzal magyarázta, hogy más volt a mérések célja és módszertana, illetve Marasek (1996) anyagában nem törekedett a szó- és mondat-hangsúlyos helyzetek szétválasztására.

A jelen tanulmány egy elővizsgálat eredményeit ismerteti, melyben három magyar magánhangzót elemeztünk két kondícióban: a legmagasabb prominenciafokon (azaz szó- és mondat-hangsúlyos pozícióban), valamint teljesen hangsúlytalan (sem szóhangsúlyt, sem mondat-hangsúlyt, de még mellékhangsúlyt sem előidéző) helyzetben, páronként azonos hangkörnyezetben és szótagszerkezetben. Azt a kérdést tettük fel, hogy a vizsgált artikulációs és akusztikai paraméterek változnak-e, és ha igen, hogyan változnak a két prominenciafok között. A szó- és mondat-hangsúly hatását tehát nem választottuk külön ebben a vizsgálatban, a két hangsúlytípus egyszerre változott (volt jelen vagy hiányzott) a két kondíció között. A vizsgálatban akusztikai paraméterekként a magánhangzók időtartamát, valamint a magánhangzók teljes időtartamán mért átlagos  $f_0$ -t, a magánhangzók középső 50 ms-án mért átlagos  $F_1$ - és  $F_2$ -értékeket, illetve ezek alapján az egyes magánhangzókhoz a magánhangzótér középpontjától számított euklideszi távolságát állapítottuk meg, míg artikulációs paraméterként a magánhangzók teljes időtartamán mért átlagos OQ értékét becsültük.

A fenti tanulmányok eredményei alapján azt vártuk, hogy a két hangsúly-kondíció között eltérnek az akusztikai és az artikulációs mérések eredményei, ugyanakkor az eltérések irányára nézve (az idézett eredmények ellentmondásai miatt) nem fogalmaztunk meg előzetes várakozásokat. Feltételeztük, hogy eltérő mintázatot tapasztalunk a nemek között az ismert fiziológiai különbségek miatt az  $f_0$  és a formánsfrekvenciák értékeiben a következők szerint. A férfiak esetében a hangszalagok hosszabbak, mint a nőknél, így a férfiak alaphangfrekvenciája jellemzően alacsonyabb, mint a nőké. Ennek következtében feltehető, hogy amennyiben a hangsúly kifejezése az  $f_0$  változtatásával jár, a változás más-más mértékű lesz/(lehet) e két beszélői csoportban, mert ebben a változásban az átlagosan használt alaphangfrekvencia meghatározó és korlátozó szerepű is egyben (úgy, hogy például a nőknél kisebb frekvenciaemelkedést tapasztalunk az eleve magasabb beszéd-alaphangfrekvenciában). Emellett általánosságban az is igaz, hogy a nők tónáléksöve jellemzően rövidebb, mint a férfiaké, így várakozásunk szerint a nőknél mért formánsfrekvenciák jellemzően magasabb frekvenciaértékeket vesznek fel, mint a férfiaknál mérték.

Ezzel a paraméterrel kapcsolatban azonban nem egyértelmű, hogy találunk-e interakciós hatást a hangsúly tényezővel. A fiziológiai eltéréseken túl a beszélő nemének interakciós hatását vártuk más, korábbi eredmények nyomán is. Ismeretes, hogy a nők általában leheletesebb zöngét produkálnak (vö. pl. Mooshammer 2010). Emellett azonban kifejezetten a zöngékepzés és a hangsúly összefüggésével kapcsolatosan például Marasek (1996) azt találta, hogy a nők és férfiak zöngéjének minősége nem változik egyező módon a hangsúly kifejeződésében sem, a nők esetében ugyanis tendenciózusan leheletesebb zöngére utaló magasabb OQ-értékeket mért, míg a férfiaknál nem talált a hangsúly és a vokális erőfeszítés között ilyen egyértelmű és rendszerszerű hatást. Mindebből kiindulva tehát interakciót vártunk a nem és a hangsúly hatásának tekintetében az  $f_0$ -ra és az OQ-ra nézve is, míg a hangsúly mellett a nem főhatását vártuk az  $F_1$ - és  $F_2$ -értékekre (valamint az ezekből származtatott távolságadatokra). A korábbi szakirodalmi eredmények alapján azt is feltételeztük, hogy a magánhangzó minőségétől függően alakulnak az időtartam, a formánsfrekvenciák és az OQ értékei a hangsúlyos és a hangsúlytalan helyzetben, de itt nem vártunk hatást (sem interakciót) a nemek szerint.

#### Anyag, kísérleti személyek és módszer

A fenti hipotézisek tesztelésére előkísérletet terveztünk. Egy korábbi, más témájú kutatáshoz (Deme et al. 2017) összeállított korpuszból válogattunk ki olyan magánhangzókat, amelyek páronként azonos hangkörnyezetben és szótagszerkezetben realizálódtak (szó- és mondat szinten is) hangsúlyos és (szó- és mondat szinten is) hangsúlytalan helyzetben (a magyar nyelv sajátosságai miatt a szóban és a mondatban elfoglalt pozíciót nem tudtuk kontrollálni, azaz a hangsúlyos helyzet mindig első szótagi, a hangsúlytalan azonban nem). Ezeknek a kritériumoknak az anyagban a felső nyelvvállású, illabiális, palatális /i/, a felső nyelvvállású, labiális, veláris /u/ és az alsó nyelvvállású, labiális, veláris /o/ felelt meg, az alábbi mondatpárokból (hs = hangsúlyos, htl = hangsúlytalan):

/o/<sub>hs</sub> *Aha. Értem.*

/o/<sub>htl</sub> *A herendi porcelánnal tálalt...*

/i/<sub>hs</sub> *Kihív.*

/i/<sub>htl</sub> *Valaki híreket hozott.*

/u/<sub>hs</sub> *Puhul a nő...*

/u/<sub>htl</sub> *A lapu hullámokban takarja el az ösvényt...*

Az /o/ magánhangzó hangsúlyos szótagbeli realizációja kapcsán felmerülhet, hogy az emocionális intonáció miatt a szó második szótagja realizálódik hangsúlyosan, ezért ellenőriztük, hogy a bemondásokban valóban az első szótag volt-e a hangsúlyos. Mivel két férfi beszélőnél ebben a szóban a vártól eltérő hangsúlymintázatot találtuk, ezekre az /o/-realizációkra vonatkozó adatokat nem elemeztük.

A kiindulásként használt korpusz felvétele során a stimulusokat (482 mondat) random sorrendben, 3-szori ismétléssel mutattuk be a SpeechRecorder (Draxler–Jänsch 2004) segítségével. A beszéd- és EGG-jelet az EGG D-200-zal (Laryngograph Ltd.) rögzítettük ECL-500M/SK csíptetős, omnidirekcionális mikrofon alkalmazásával, 44,1 kHz-es mintavételi frekvencián, csendesített szobában. A jelen vizsgálatban 18 beszélő (7 nő, 11 férfi) felolvasását elemeztük, akiknek az életkora 23 és 43 év között volt (átlag = 32,2 év, szórás = 3,8 év). Egyiküknek sem volt sem beszéd-, sem hallásproblémája. (Az eredeti korpusz egyik női beszélőjének az anyagát az EGG technikai problémája miatt ki kellett hagynunk az elemzésből, továbbá ahogyan azt már említettük, kizártuk két férfi néhány /v/-realizációját.) Mindösszesen 101 /v/-, 108 /i/- és 108 /u/-megvalósulást elemeztünk.

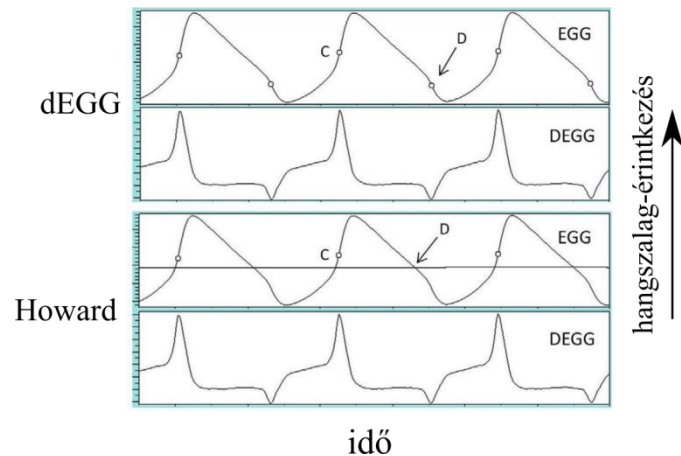
A felvételekben a célmagánhangzókat manuálisan címkéztük a Praat szoftverben (Boersma–Weenink 2016). A célmagánhangzókon a következő paramétereket mértük meg: (i) a magánhangzó időtartama, (ii) az  $F_1$  és  $F_2$  átlagos értékei a magánhangzó középső 50 ms-os szegmensében, (iii) az  $f_0$  átlaga a teljes magánhangzó-időtartamban az EGG-jel alapján, (iv) az OQ átlaga a teljes magánhangzó-időtartamban az EGG-jel alapján. A magánhangzó időtartamát és a formánsértékeket a Praat szoftver segítségével mértük, és az utóbbiakat az ábrázoláshoz  $z$ -transzformációval (Lobanov 1971) beszélőnként standardizáltuk az R szoftverben (R Core Team 2017) a phonR csomag használatával (McCloy 2016). Emellett kiszámítottuk az egyes magánhangzóknak a magánhangzó-tér közepétől mért euklideszi távolságát ( $|d|$ ) is a formánsfrekvenciák alapján (Bradlow et al. 1996), és ezt a paramétert is elemeztük függő változóként.

Az EGG-jel elemzését a Peakdet Matlab-szkript (Michaud–Thi 2007) Praat szkriptnyelven implementált változatával (Kirby 2017) végeztük, mely az eredeti szkripttel ellentétben a derivált jel csúcsait autokorreláció (Boersma 1993) segítségével detektálja. A férfiak esetében 70–250 Hz, a nők esetében 100–350 Hz közötti  $f_0$ -tartomány megadásával dolgoztunk. A jelek parametrizálását megelőzően mind az EGG-, mind a dEGG-jelet 10 lépéses mozgóátlag-szűrővel simítottuk.

Az OQ és az  $f_0$  becsléséhez a hangszalagok nyitódásának és záródásának időpontját kétféle módszerrel is meghatároztuk (1. ábra). A záródási pontnak (az ábrán C betűvel jelölve) és a nyitódási pontnak (az ábrán D betűvel jelölve) a két módszer által meghatározott helyzetét az 1. ábra illusztrálja az eredeti és a derivált jel megjelenítésével. Egyrészt alkalmaztuk a dEGG-módszert (Henrich et al. 2004), amely az EGG-jel első deriváltjának legnagyobb pozitív és legalacsonyabb negatív pontjában jelöli meg a záródás, illetve a nyitódás pillanatát (a továbbiakban  $OQ_{dEGG}$ ). Másrészt egy kevert módszert is használtunk, melyet Howard (1995) dolgozott ki (a továbbiakban:  $OQ_{Howard}$ ). A záródás pillanatát ez a módszer is az első derivált maximumánál határozza meg, azonban a nyitódást arra a pillanatra teszi, ahol az

eredeti jel negatív irányba haladva átlépi az eredeti EGG-jel lokális maximuma és minimuma közötti 3/7-es amplitúdóküszöböt. Az így kapott nyitódási és záródási időpontok alapján az egymást követő záródási pontok által határoljuk el a periódusokat (egy periódus tehát a két szomszédos C pont által határolt szakasz). Ennek megfelelően egy periódusra az  $f_0$  értéke a periódust határoló záródási pontok időbeli távolságának (periódusidő) a reciproka, míg az OQ értékét úgy kapjuk meg, hogy a periódus kezdete és a nyitódási pont közötti időkülönbséget elosztjuk a periódusidővel.

Amennyiben a két becslési módszer esetén az egy magánhangzóra kapott OQ-értékek jelentősen eltértek – tekintve, hogy ez a csúcsok téves detektálásának jó indikátora –, az adott eseteket ellenőriztük, és a detektált dEGG-csúcsok helyzetét manuálisan korrigáltuk.



1. ábra

A hangszalagok záródásának (C-vel jelölt pont) és nyitódásának (D-vel jelölt pont) meghatározása a dEGG-módszer (fent) és Howard hibrid módszere (lent) esetén (Awan és Awan 2013 Figure 2 alapján, 436)

#### Statisztikai elemzések

Az R szoftverben (R Core Team 2017) lineáris kevert modellekkel (Kuznetsova et al. 2017) elemeztük a hangsúlyosság hatását az időtartamra, az első két formánsra, a magánhangzóknak a magánhangzótér középpontjától mért euklideszi távolságára, az alapfrekvenciára és a kétféle OQ-értékre. Az említett fiziológiai okokból következően amennyiben a függő változó a fonáció valamely jellemzőjét ragadta meg ( $f_0$ ,  $OQ_{dEGG}$ ,  $OQ_{Howard}$ ), vagy a toldalékcso tulajdonságait képezte le ( $F_1$ ,  $F_2$ ,  $|d|$ ), a modellek a hangsúlyosság és a magánhangzó-minőség mellett független változóként tartalmazták az adatközlő nemét, illetve a nem, a magánhangzó-minőség és a hangsúlyosság in-



terakcióját. A modellek random hatásként az adatközlőt tartalmazták, adatközlőnként random meredekséggel. A variancia homogenitását és a normalitást a reziduálisok grafikus elemzése alapján állapítottuk meg. A  $p$ -értékek számítása Satterthwaite-approximációval történt. A modelleken Tukey-féle post hoc teszttel végeztünk páronkénti összevetéseket. Mindezek mellett a két eltérő OQ-mérés eredményeit is összevetettük Pearson-féle korrelációelemzéssel, hogy az eltéréseiket feltérképezhessük.

### Eredmények

A kapott adatok átlag- és szórásértékeit nemenként, magánhangzónként és hangsúlyhelyzetenként az 1. táblázat összegzi.

1. táblázat: A hangsúlyos és a hangsúlytalan helyzetű magánhangzókra kapott mérési eredmények nemenként és magánhangzónként (átlag és szórás)

Nők						
	/ɒ/		/i/		/u/	
	Hs	Htl	Hs	Htl	Hs	Htl
<b>Időtartam (ms)</b>	77±27	45±18	67±18	54±12	70±14	73±17
<b>F<sub>1</sub> (Hz)</b>	774±123	807±121	390±63	400±61	430±45	434±48
<b>F<sub>2</sub> (Hz)</b>	1198±137	1394±220	2024±589	2108±597	863±358	882±354
<b>f<sub>0</sub> (Hz)</b>	200±39	207±36	192±32	211±33	241±54	248±39
<b>OQ<sub>dEGG</sub></b>	0,66±0,07	0,64±0,09	0,60±0,06	0,61±0,06	0,66±0,05	0,63±0,04
<b>OQ<sub>Howard</sub></b>	0,62±0,07	0,61±0,08	0,58±0,06	0,59±0,06	0,61±0,06	0,60±0,05

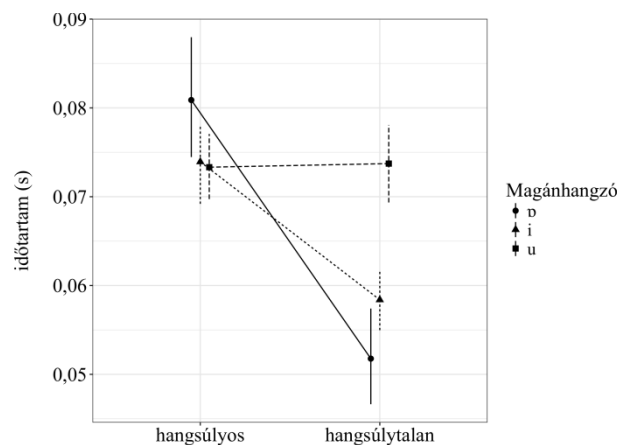
Férfiak						
	/ɒ/		/i/		/u/	
	Hs	Htl	Hs	Htl	Hs	Htl
<b>Időtartam (ms)</b>	83±23	57±20	78±15	61±13	75±14	74±14
<b>F<sub>1</sub> (Hz)</b>	642±82	640±151	337±38	319±28	354±35	376±51
<b>F<sub>2</sub> (Hz)</b>	1068±148	1288±378	2205±229	2206±255	863±313	866±445
<b>f<sub>0</sub> (Hz)</b>	118±17	124±18	129±23	125±20	145±27	162±25
<b>OQ<sub>dEGG</sub></b>	0,58±0,07	0,68±0,08	0,55±0,08	0,57±0,07	0,58±0,07	0,57±0,08
<b>OQ<sub>Howard</sub></b>	0,61±0,06	0,66±0,06	0,56±0,05	0,60±0,07	0,58±0,05	0,57±0,08

A továbbiakban az egyes paraméterek jellemzésekor a hangsúlytalan és a hangsúlyos helyzet között mért eltéréseket számszerűsítjük, először a teljes adathalmazra vonatkoztatva, majd külön nemenként és magánhangzóminőségenként is. Az adatok bemutatásában a könnyebb áttekinthetőség kedvéért azt az eljárást alkalmazzuk, hogy a hangsúlytalan realizációk átlagértékét tekintjük a kiindulásnak, ehhez viszonyítjuk a hangsúlyos realizációk átlagértékét. A hangsúlyos ejtés magasabb értéke esetén a különbséget pozitív

előjellel számszerűsítjük, míg ha a hangsúlytalan szótagbeli ejtés értéke magasabb, a különbséget negatív előjellel adjuk meg. Példaképpen: a hangsúlytalan ejtésben mért valamely értéket tekintjük 100%-nak, így ha a hangsúlyos értéke 105%, azt +5%-ként, ha hangsúlyosé 95%, azt -5%-ként tüntetjük fel.

### A magánhangzók időtartama

A magánhangzók időtartamát a hangsúly és a magánhangzó-minőség függvényében elemeztük (2. ábra). Szignifikáns interakciót találtunk a hangsúly és a magánhangzó-minőség között [ $F(2, 282,85) = 25,015$ ;  $p < 0,001$ ]: míg mind az /o/, mind az /i/ esetében szignifikánsan ( $p < 0,001$ ) hosszabbak voltak a hangsúlyos magánhangzók, az /u/ esetében a hangsúly hatása nem volt szignifikáns, mind a hangsúlyos, mind a hangsúlytalan /u/ az /i/ hangsúlyos realizációjára jellemző időtartammal valósult meg. Az egyes magánhangzók időtartam-eltérése hasonlóan alakult a két nemnél: az /o/ a nők ejtésében +32,1 ms (+71,8%), a férfiakéban +26,5 ms (+46,7%); az /i/ a nőknél +13,0 ms (+24,0%) és a férfiaknál +17,2 ms (+28,1%); az /u/ a nőknél -2,7 ms (-3,7%) és a férfiaknál +0,9 ms (+1,2%) eltérést mutatott a hangsúlyos és hangsúlytalan előfordulások között.



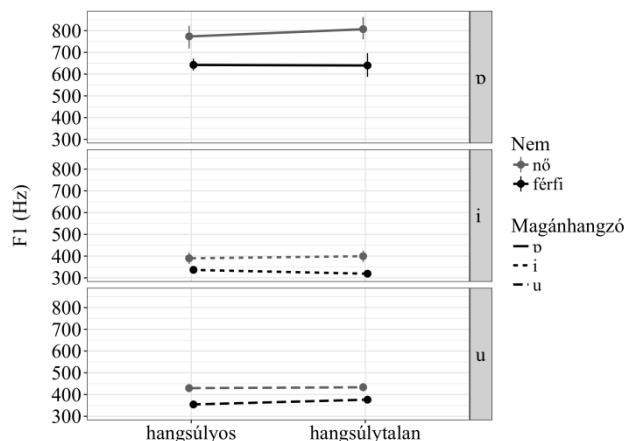
2. ábra

A magánhangzók időtartama (átlag és 95%-os konfidenciaintervallum) a hangsúlyviszonyok függvényében

### Formánsértékek

A formánsértékeket a hangsúly, a magánhangzó-minőség és a nem függvényében elemeztük (3. ábra). A statisztikai elemzés szerint a hangsúlyosság önmagában nem bizonyult meghatározónak az első formáns értékeinek alaku-

lására [ $F(1, 98,74) = 0,91, p = 0,342$ ], ugyanis sem az összes beszélő együttes adataiban ( $-5,5$  Hz, azaz  $-1,1\%$ ), sem a női ( $-15,4$  Hz különbség, azaz  $-2,8\%$ ), illetve a férfi adatközlők ( $+5,2$  Hz különbség, azaz  $+1,2\%$ ) csoportjaiban, sem magánhangzó-minőségre lebontva nem volt kimutatható eltérés a magánhangzók  $F_1$ -értékében a hangsúlyos és a hangsúlytalan helyzetek között. Az /i/ megvalósulásaiban a nők esetében átlagosan  $-9,8$  Hz, azaz  $-2,4\%$  az első formáns különbsége, a férfiaknál  $+17,6$  Hz, azaz  $+5,5\%$ . Az /v/-ban a nőknél  $-33,4$  Hz ( $-4,1\%$ ), a férfiaknál  $+2,5$  Hz ( $+0,4\%$ ) az első formáns eltérése. Az /u/ megvalósulásaiban az  $F_1$   $-4,0$  Hz ( $-0,9\%$ ) különbséget mutatott a női és  $-21,9$  Hz-et ( $-5,8\%$ ) a férfiejtésben a hangsúlyosságtól függően.

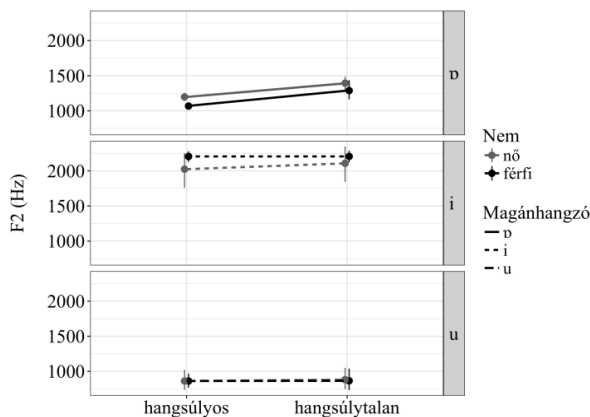


3. ábra

Az első formáns értékei (átlag és 95%-os konfidenciaintervallum) a vizsgált tényezők függvényében

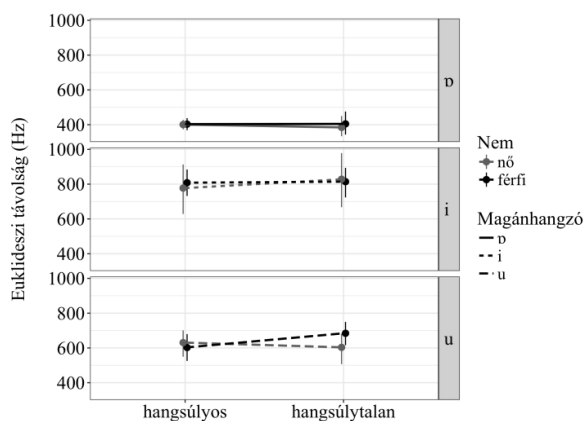
Az  $F_2$  tekintetében szignifikáns eltérés volt kimutatható a hangsúlyhelyzetek között [ $F(1, 79,448) = 4,88; p = 0,030$ ] (4. ábra): összességében a hangsúlyos magánhangzók második formánsai alacsonyabb értékeket vettek fel ( $-110,0$  Hz, azaz  $-7,5\%$ ). A páronkénti összevetés szerint azonban az adatokat nemenként és magánhangzó-minőségenként szétbontva nem voltak szignifikáns különbségek a hangsúlyos és hangsúlytalan ejtés között, az eltérés elsősorban minden bizonnyal az /v/ magánhangzónak köszönhető: az alsó nyelvvállású /v/ esetében a nők ejtésében  $-196,8$  Hz ( $-14,1\%$ ), a férfiakéban  $-220,3$  Hz ( $-17,1\%$ ) volt az  $F_2$  különbsége a két hangsúlyhelyzetben. Az /u/-ra a nőknél  $-19,7$  Hz ( $-2,2\%$ ), a férfiaknál  $-2,3$  Hz ( $-0,3\%$ ); az /i/-re a nőknél  $-83,3$  Hz ( $-4,0\%$ ), a férfiaknál  $-0,4$  Hz ( $-0,02\%$ ) különbséget kaptunk a hangsúlyosság függvényében.

A magánhangzó-realizációknak a magánhangzótér középpontjától számított euklideszi távolságát nem befolyásolta a prominencia, erre a változóra csak a magánhangzó-minőség volt szignifikáns hatással [ $F(2, 300,19) = 97,66$ ;  $p < 0,001$ ] (5. ábra), ami azt jelenti, hogy az egyes magánhangzók tendenciózusan eltérő távolságra realizálódtak az akusztikai magánhangzótér középpontjától, függetlenül attól, hogy hangsúlyos vagy hangsúlytalan szótagban realizálódtak-e.



4. ábra

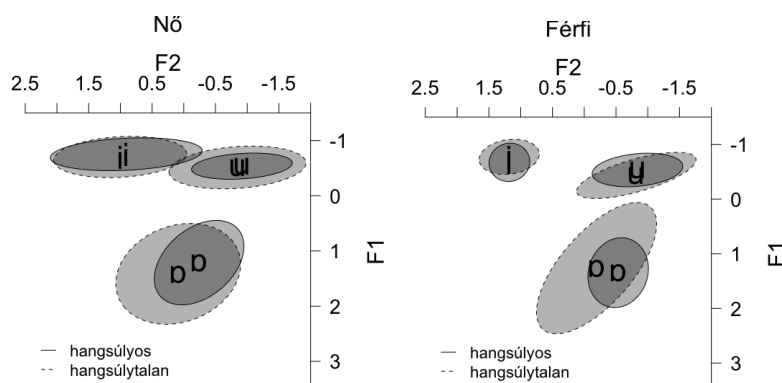
A második formáns értékei (átlag és 95%-os konfidenciaintervallum) a vizsgált tényezők függvényében



5. ábra

A magánhangzó-tér középpontjától számított távolság (átlag és 95%-os konfidenciaintervallum) a vizsgált tényezők függvényében

Annak ellenére, hogy csak a második formáns értékeiben találtunk szignifikáns eltérést a hangsúlyosság mentén, a 6. ábrán látható sűrűsödési ellipszisekről leolvasható, hogy a formánsok frekvenciája és a hangsúly között általánosságban is van összefüggés, ugyanis a frekvenciaértékek variabilitása nagyobb volt a hangsúlytalan pozícióban, legalábbis a hátul képzett magánhangzók esetében.



6. ábra

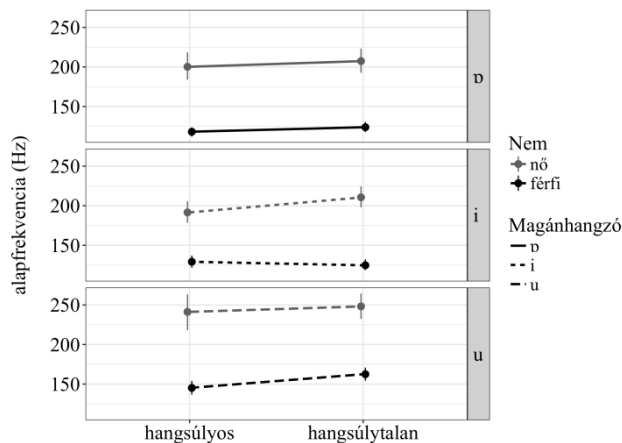
Az elemzett magánhangzók z-transzformált  $F_1$ - és  $F_2$ -alapú sűrűsödési ellipszisei (átlag  $\pm$  szórás mindkét dimenzió mentén)

A z-transzformált értékekre a hangsúlytalan /u/  $F_1$ -ének szórása 0,258, az  $F_2$  szórása 0,623; a hangsúlyos /u/  $F_1$ -ének a szórása 0,178, az  $F_2$ -é 0,468 volt. A hangsúlytalan /v/  $F_1$ -ének szórása 0,684, az  $F_2$ -é 0,614; a hangsúlyos /v/  $F_1$ -értékének a szórása 0,435, az  $F_2$  szórása 0,384 volt. Az /i/ esetében nem tapasztaltunk ilyen mintázatot: a hangsúlytalan megvalósulások  $F_1$ -értékeinek szórása 0,217, az  $F_2$  szórása 0,473, hangsúlyos helyzetben pedig 0,211, illetve 0,504 volt az első, illetve a második formáns standardizált értékeinek szórása. A tendencia mindkét nem esetében érvényesült mindkét veláris magánhangzóban: az /v/ esetén a nőknél hangsúlyos helyzetben 0,476, illetve 0,440, hangsúlytalan helyzetben pedig 0,568, illetve 0,612 volt az első, illetve a második formáns z-transzformált értékeinek szórása. A férfiaknál hangsúlyos helyzetben 0,407, illetve 0,303, a hangsúlytalan magánhangzók-nál 0,760, illetve 0,598 voltak a szórásértékek. Az /u/ esetében a nőknél hangsúlyos helyzetben 0,148, illetve 0,496, hangsúlytalan helyzetben 0,241, illetve 0,674 formáns-szórásértékeket mértünk, a férfiaknál hangsúlyos helyzetben 0,196, illetve 0,456, hangsúlytalan helyzetben 0,268, illetve 0,596 volt az  $F_1$ , illetve az  $F_2$  értékeinek szórása. Az /i/ megvalósulásai esetén csak a férfiak második, illetve a nők első formánsa követte ezt a tendenciát: míg a nőknél hangsúlyos helyzetben 0,186, illetve 0,747, hangsúlytalan helyzetben

0,233, illetve 0,653 volt a két formáns szórásértéke, a férfiaknál hangsúlyos helyzetben 0,223, illetve 0,204, hangsúlytalan ejtésben 0,205, illetve 0,301 volt a formánsértékek szórása.

#### A magánhangzók alapfrekvenciája

A vizsgált magánhangzókban mért átlagos alapfrekvencia a nőknél 217 Hz (szórás 44 Hz), a férfiaknál 134 Hz (szórás 27 Hz) volt. Az alapfrekvencia variabilitásának statisztikai elemzését a hangsúly, a magánhangzó-minőség és a nem függvényében végeztük el (7. ábra).



7. ábra

Az alapfrekvencia értékei (átlag és 95%-os konfidenciaintervallum) a vizsgált tényezők függvényében

Hármas interakciót találtunk a hangsúly, a nem és magánhangzó-minőség független változók között [ $F(2, 281,14) = 5,162$ ;  $p = 0,006$ ]. A páronkénti összevetés alapján egyik magánhangzó-minőségnél és egyik nemnél sem volt szignifikáns különbség a hangsúlyos és a hangsúlytalan helyzetben mért értékek között. A szignifikáns interakció oka feltehetőleg az, hogy az /i/ ejtésében nemenként eltérő tendencia bontakozik ki: bár egyik nemnél sem szignifikáns az eltérés, a férfiak hangsúlyos helyzetben ejtenek magasabb alapfrekvenciájú magánhangzókat (a különbség +4,4 Hz, azaz +3,6%), míg a nők esetében alacsonyabb a hangsúlyos /i/  $f_0$ -ja: az eltérés -19,1 Hz (-9,1%). Az /v/ és az /u/ esetében mindkét nem ejtésében a hangsúlyos helyzetben mérünk alacsonyabb  $f_0$ -értéket [/v/ nők: -7,2 Hz (-3,5%), férfiak: -5,7 Hz (-4,6%); /u/ nők: -7,0 Hz (-2,8%), férfiak -17,3 Hz (-10,7%)]. Ezek a kis-mértékű eltérések nem bizonyultak szignifikánsnak a prominencia jelölésében.

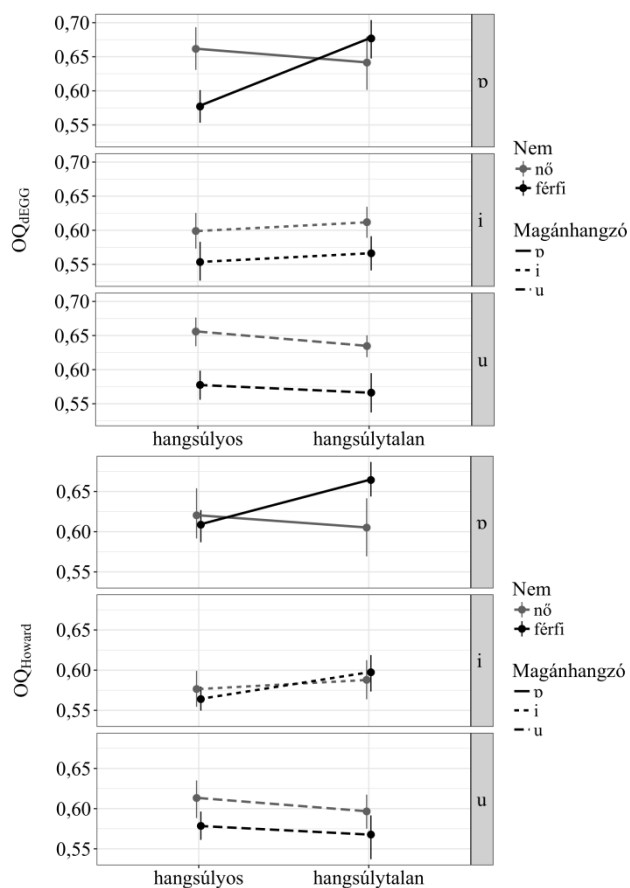
### Az OQ-értékek

Az OQ-értékeket a hangsúly, a magánhangzó-minőség és a nem függvényében elemeztük (8. ábra). Az  $OQ_{dEGG}$  átlagos értékei nőknél 0,63 (szórás: 0,07), férfiaknál 0,58 (szórás: 0,08) voltak. Az  $OQ_{Howard}$  átlagos értéke mind a nők, mind a férfiak esetében 0,60 volt (a szórás értéke a nőknél 0,06, a férfiaknál 0,07). A két becslési módszer közül tehát a dEGG pozitív és negatív csúcsain alapuló módszer volt érzékenyebb olyan jellegű különbségekre, amelyek a nők és férfiak csoportját a jelen adathalmazban elkülönítik. A továbbiakban ezért a két mérési móddal kapott eredményeket külön közöljük.

Az  $OQ_{dEGG}$ -értékek esetében háromszoros interakció volt megfigyelhető a hangsúly, a nem és a magánhangzó-minőség között [ $F(2, 282,26) = 6,959$ ;  $p = 0,001$ ]. A Tukey-féle post hoc teszt alapján az /v/ esetében a férfiaknál szignifikánsan alacsonyabb OQ-val valósultak meg a hangsúlyos magánhangzók, mint a hangsúlytalanok ( $p < 0,001$ ): az átlagok eltérése  $-0,10$  ( $-14,7\%$ ). A nőknél ezzel szemben  $+0,02$  ( $+3,1\%$ ) magasabb  $OQ_{dEGG}$ -értéket mértünk a hangsúlyos helyzetben, bár nem volt szignifikáns a különbség. A másik két magánhangzó-minőség esetében nem tapasztaltunk szignifikáns eltérést, az OQ-értékek átlagai az /i/ esetében valamennyivel alacsonyabbak, az /u/ esetében valamennyivel magasabbak voltak hangsúlyos helyzetben, mint hangsúlytalanban: az /i/-ben a nőknél  $-0,01$  ( $-1,6\%$ ), a férfiaknál  $-0,02$  ( $-3,5\%$ ); az /u/-ban a nőknél  $+0,03$  ( $+4,8\%$ ), a férfiaknál  $+0,01$  ( $+1,8\%$ ).

Az  $OQ_{Howard}$ -mérésekre kapott eredmények többnyire az  $OQ_{dEGG}$ -adatokhoz hasonlóan alakultak, de eltéréseket is találtunk. A hangsúly hatásával külön-külön interakcióban volt a nem [ $F(1, 17,96) = 5,799$ ;  $p = 0,027$ ] és a magánhangzó-minőség [ $F(2, 281,69) = 3,889$ ;  $p = 0,022$ ] is. Az  $OQ_{dEGG}$  esetében tapasztalt interakcióhoz hasonlóan az /v/ ejtésében a férfiaknál ismét szignifikánsan alacsonyabb [ $-0,05$  ( $-7,6\%$ )] volt az OQ értéke hangsúlyos helyzetben ( $p = 0,011$ ), míg a nőknél nem szignifikáns, de ellentétes irányú eltérés [ $+0,01$  ( $+1,6\%$ )] volt megfigyelhető. A másik két magánhangzónál ismét nem mutatott ki szignifikáns különbségeket a páronkénti összevetés a hangsúly függvényében, azonban az átlagértékek eltérésének iránya azonos volt a másik mérési módszernél tapasztaltakéval: az /i/ esetében az  $OQ_{Howard}$ -átlagértékek alacsonyabbak voltak hangsúlyos helyzetben, míg az /u/ esetében a hangsúlytalanoknál magasabb értékeket mutattak a hangsúlyos magánhangzók nemtől függetlenül: az /i/-ben a nőknél  $-0,01$  ( $-1,7\%$ ), a férfiaknál  $-0,04$  ( $-6,7\%$ ); az /u/-ban a nőknél  $+0,01$  ( $+1,7\%$ ), a férfiaknál  $+0,01$  ( $+1,8\%$ ).

A két mérési módszer (amelyeknek az eredményeit az 8. ábrán mutatjuk be) között tehát valóban volt eltérés az adatok mintázatában. Ugyanakkor az azonos magánhangzó-minőségekre kapott mérési eredményeket összevetve erős korrelációt kaptunk a beszélő neme, a magánhangzó-minőség és a prominencia kontrollálása mellett a Pearson-próbával ( $r = 0,745$ ;  $p < 0,001$ ).



8. ábra

Az OQ mérésére alkalmazott módszerekkel kapott eredmények (átlag és 95%-os konfidenciaintervallum)

### Következtetések

A jelen tanulmányban három magyar magánhangzót elemeztünk a prominencia függvényében. A magánhangzókban az időtartam, az alaphfrekvencia, az első és második formáns, valamint az OQ értékeit mértük (az utóbbit kétféle mérőmódszertannal is) és vetettük össze kétféle hangsúlykondícióban: szó- és mondathangsúlyos helyzetben és teljesen hangsúlytalan pozícióban. Az OQ tekintetében a magyar beszédre vonatkozóan elsőként közöltünk adatokat.



A magánhangzók időtartama jelentősen eltért a hangsúlyosság függvényében: a (szó- és mondat)hangsúlyos magánhangzók szignifikánsan hosszabbak voltak, mint a hangsúlytalan realizációk, hasonlóan Szalontai és munkatársai (2016), illetve Mády és kollégái (2017) eredményeihez, akik mind a szóhangsúly, mind a mondathangsúly kifejeződésében kimutatták az időtartam szerepét. Magdics (1966) közleményében a tendencia hasonló: a hangsúlyos helyzetű magánhangzókra hosszabb időtartamokat mért (bár statisztikai összevetést nem végzett); Kovács (2002) ugyanakkor szignifikánsan rövidebb időtartamokat adatolt a hangsúlyos pozícióban. Kassai (1979) és Gósy (1997) munkáiban nem látszanak egységes tendenciák az időtartam és a hangsúlyosság összefüggésében, aminek számos oka lehet, köztük az, hogy a szerzők az említett vizsgálatok mindegyikében merőben eltérő anyagot és módszertanokat alkalmaztak, és jelentősen eltérő számú beszélő felvételeit elemezték.

Az alapfrekvenciára kapott eredményeink eltérnek Szalontai és munkatársai (2016), illetve Mády és kollégái (2017) eredményeitől. Ennek oka ugyan- csak a módszertan és a vizsgálati anyag különbsége lehet.

A formánsadatokat kétféle módon elemeztük. Egyrészt magukon az  $F_1$  és  $F_2$  formánsértékeken is végeztünk statisztikai elemzéseket. Másrészt ezek alapján meghatároztuk az egyes magánhangzó-megvalósulásoknak a magánhangzóter közép pontjától számított euklideszi távolságát, és ezeket vetettük össze a prominencia függvényében. Az euklideszi távolságok és az  $F_1$  nem mutattak összefüggést a hangsúlyhelyzettel, kizárólag az  $F_2$ -értékek esetében találtunk ilyen hatást. Itt ismét utalhatunk Lindblom (1963) azon feltételezésére, amely szerint a magánhangzó-realizáció időtartama és formánsértékei között összefüggés volna, oly módon, hogy a rövidebb magánhangzókban célalulmúlás történik (mivel a beszédsszerveknek „nincs idejük” elérni a célkonfigurációt), míg a hosszabb időtartamban realizálódó beszédhangok esetén a célkonfiguráció megvalósul. Bár statisztikai eltérést nem találtunk, az jól látszik, hogy a hangsúlyos (és egyúttal hosszabb időtartamban realizált) magánhangzók formánsértékei kisebb mértékben szóródtak, mint a hangsúlytalan (és egyúttal rövidebb) magánhangzókéi (vö. 6. ábra). Azaz bár a vizsgált anyagban átlagosan valóban nem mutatkozott különbség a magánhangzók ejtése között a centralizáltság tekintetében, a hosszabb, azaz a hangsúlyos magánhangzók esetében nagyobb arányban fordultak elő periferikusabb ejtésűek – legalábbis a formánsértékek alapján – ami egybecseng Lindblom (1963) feltételezésével.

Ha a Magdics (1965) által nemek és hangsúlyhelyzetek szerint elkülönítve megadott átlagos formánsértékeket részleteiben is összevetjük a saját megfelelő adatainkkal, a következőket találjuk. Magdics női adatközlőinél mindkét formáns mindhárom, általunk is elemzett magánhangzó esetében magasabb átlagos frekvencia-középpértékkel bírt a hangsúlyos helyzetben (az eltérések az  $F_1$ -ben +6,0 és +13,1% közöttiek, az  $F_2$ -ben +7,4 és +8,0% közöttiek), míg a jelen kutatás női adatközlőinek ejtésében minden esetben a hangsúlytalan

helyzetben mértünk átlagosan magasabb formánsfrekvenciákat (az eltérések az F<sub>1</sub>-ben -0,9 és -4,1% közöttiek, míg az F<sub>2</sub>-ben -2,2 és -14,1% között vannak). A férfiak esetében Magdics ugyancsak mindenhol átlagosan magasabb formánsfrekvenciákat adott meg a hangsúlyos szótagok magánhangzóira (az eltérések az F<sub>1</sub>-ben +12,5 és +21,1% közöttiek, az F<sub>2</sub>-ben +3,2 és +10,6% közöttiek), a jelen elemzés férfi adatközlőinek ejtésében azonban még az eltérések iránya sem egységes (az F<sub>1</sub>-ben -5,9 és +5,6% közöttiek, míg az F<sub>2</sub>-ben 0,0 és -17,1% között vannak). Érdekes megjegyezni, hogy az anyag és a módszer eltérései okozhatják ezt a divergenciát az adatokban, ami ugyanakkor utalhat arra is, hogy a formánsfrekvenciák nem akusztikai korrelátumai a prominenciának a magyarban. Ezzel együtt – mint láttuk – a megvalósítás variabilitása összefüggeni látszik a hangsúlyhelyzettel.

Az OQ kétféle, itt alkalmazott becslési módszere közül csak a dEGG pozitív és negatív csúcsain alapuló módszer mutatott ki a nemek között különbségeket. Ugyanezt az eltérést találta az EGG-jel különböző parametrizálási módjainak összevetésekor Awan és Awan (2013) 25 fiatal nő és 25 fiatal férfi adatai alapján. A módszerek közötti különbséget a szerzők azzal magyarázzák, hogy a hangszalagok gyors szétválását, mely inkább a férfiakra jellemző, a küszöbértékeken alapuló módszerek nem képesek detektálni, hanem még a szétválás valódi pillanata előttre teszik azt, így a valódinál rövidebb hangszalag-érintkezéssel (és ennek következtében magasabb OQ-val) jellemzik a periódust. Ez az eltérés jól látszik az 1. ábrán.

A jelen kutatás eredményei azt mutatták, hogy a prominencia ugyan hat az OQ-értékekre, de ez a hatás nem általános, hanem a nemmel és magánhangzó-minőséggel interakcióban látszik. Elsősorban ugyanis a férfiak beszédében, és azon belül is csak az /v/ magánhangzóval kapcsolatban figyeltük meg. Az /v/ esetében a férfiaknál mindkét OQ-számítási metódus esetén nagyobb negatív irányú eltérést mértünk a hangsúlyosság hatására az OQ-értékekben, ami a nagyobb vokális erőfeszítéssel függhet össze; míg a nőknél ennél jóval kisebb, (nem szignifikáns) pozitív irányú eltérést adatoltunk a hangsúlytalan helyzethez képest, ami a hangsúlyos helyzetű magánhangzók esetében kisebb mértékben levegősebb zöngéképzésre utalhat. Ugyan Marasek (1996) is talált interakciót a nem és a hangsúly hatása között a német beszédén végzett vizsgálata során, azonban az ő adataiban a nők mutattak nagyobb (bár a jelen vizsgálatban találtakhoz hasonlóan pozitív irányú) eltérést, a férfiaknál pedig nem tapasztalt különbséget a két kondíció között.

Összegezve: részben a korábbi kutatásokkal egybevetően azt találtuk, hogy a magánhangzó-időtartam akusztikai korrelátuma a mondathangsúlynak a magyarban – a magánhangzó-minőségtől és a beszélő nemétől függetlenül –, a jelen vizsgálatban azonban a további akusztikai és artikulációs paraméterek elemzése nem mutatott a hangsúlyhelyzettől függő egyértelmű eltéréseket. Mindennek ellenére például az OQ artikulációs paraméterrel kapcsolatosan olyan tendenciákat figyelhetünk meg, amelyek további vizsgálatát követ-

lenül érdemesnek és relevánsnak tartjuk. Ez az eredmény pedig a mérőszám-  
nak a magánhangzó-minőség és a nem függvényében felsejlő eltérése, amely  
hatások megléte okvetlenül tisztázandó kérdés. Bár az OQ mint artikulációs  
paraméter nem mutatta a hangsúlyhelyzet hatását, úgy gondoljuk, hogy  
Fónagynak (1958) az a feltételezése, amely szerint az akusztikai paraméterek  
helyett az artikulációs elemzések vihetnek közelebb a hangsúly megvalósulá-  
sának módzataihoz, nem vethető el, és e felvetés ellenőrzésére további arti-  
kulációs (például ultrahanggal vagy elektromágneses artikulográffal végzett)  
vizsgálatok szükségesek. Ezek az elemzések pedig remélhetőleg árnyalják a  
jelen elemzésben kapott képet, valamint tisztázzák a prominencia artikulációs  
implementációjának lehetséges kérdéseit is.

### Irodalom

- Awan, Shaheen N. – Awan, Jordan A. 2013. The effect of gender on measures of  
electroglottographic contact quotient. *Journal of Voice* 27/4. 433–440.
- Blaho, Sylvia – Szeredi, Dániel 2011. Secondary stress in Hungarian: (Morpho)-  
syntactic, not metrical. In Washburn, Mary Byram – McKinney-Bock, Katherine –  
Varis, Erika – Sawyer, Ann – Tomaszewicz, Barbara (eds.): *Proceedings of the  
28th West Coast Conference on Formal Linguistics*. Cascadia Proceedings Pro-  
ject, Somerville, MA. 51–59.
- Bradlow, Ann M. – Torretta, Gina M. – Pisoni, David B. 1996. Intelligibility of nor-  
mal speech I: Global and fine-grained acoustic-phonetic talker characteristics.  
*Speech Communication* 20. 255–272.
- Boersma, Paul 1993. Accurate short-term analysis of the fundamental frequency and  
the harmonics-to-noise ratio of a sampled sound. In: *Proceedings of the Institute of  
Phonetic Sciences* 17. University of Amsterdam, Amsterdam. 97–110.
- Boersma, Paul – Weenink, David 2016. Praat: doing phonetics by computer.  
<http://www.fon.hum.uva.nl/praat/>. (A letöltés ideje: 2016. november 4.)
- Childers, Donald G. – Lee, Chih K. 1990. Vocal quality factors: Analysis, synthesis,  
and perception. *Journal of the Acoustical Society of America* 90/5. 2394–2410.
- Deme, Andrea – Bartók, Márton – Grácz T. E., Tekla Etelka – Markó, Alexandra –  
Varjasi, Gergely – Csapó, Tamás Gábor 2017. *Intervocalic voicing of the Hungari-  
an /h/*. Előadás. 13<sup>th</sup> International Conference on the Structure of Hungarian, Buda-  
pest, 2017. június 29–30.
- Draxler, Christoph – Jänsch, Klaus 2004. SpeechRecorder – a universal platform in-  
dependent multi-channel audio recording software. In: *Proceedings of International  
Conference on Language Resources and Evaluation*. Lisbon. 559–562.
- Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In Hardcastle, Wil-  
liam J. – Laver, John – Gibbon, Fiona E. (eds.): *The handbook of phonetic sciences*.  
Second edition. Blackwell, Oxford. 521–602.
- Fónagy Iván 1958. *A hangsúlyról*. Nyelvtudományi Értekezések 18. Akadémiai Ki-  
adó, Budapest.
- Genzel, Susanne – Ishihara, Shinichiro – Surányi, Balázs 2015. The prosodic expres-  
sion of focus, contrast and givenness: A production study of Hungarian. *Lingua*  
165. Part B. 183–204.

- Gósy Mária 1997. Semleges magánhangzók a magyar beszédben. *Magyar Nyelvőr* 121. 9–19.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134/2. 140–165.
- Henrich, Nathalie – d’Alessandro, Christophe – Doval, Boris – Castellengo, Michèle 2004. On the use of the derivative of electroglottographic signals for characterization of nonpathological voice phonation. *Journal of the Acoustical Society of America* 115. 1321–1332.
- Howard, David M. 1995. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice* 9/2. 163–172.
- Imrényi András 2007. A magyar szórend kísérleti modelljei II. Tartományok és viszonyok a magyar mondatban. *Magyar Nyelvőr* 131/4. 430–451.
- Kálmán László – Nádasdy Ádám 1994/2016. A hangsúly. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 393–467. (Digitális kiadás: 2016.)
- Kankare, Elina – Laukkanen, Anne-Maria – Ilomäki, Irma – Miettinen, Anne – Tiina Pyllkkänen 2012. Electroglottographic contact quotient in different phonation types using different amplitude threshold levels. *Logopedics Phoniatrics Vocology* 37/3. 127–132.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.
- Keszler Borbála szerk. 2000. *Magyar grammatika*. Nemzeti Tankönyvkiadó, Budapest.
- Kirby, James 2017. *Praatdet: Praat-based tools for EGG analysis* (v. 0.1.1). <https://doi.org/10.5281/zenodo.1117189>.
- Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. PhD-értekezés. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen. <https://dea.lib.unideb.hu/dea/bitstream/handle/2437/79756/ertekezes.pdf?sequence=6&isAllowed=y>. (A letöltés ideje: 2018. január 27.)
- Kuznetsova, Alexandra – Brockhoff, Per B. – Christensen, Rune 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82/13. 1–26. <http://doi.org/10.18637/jss.v082.i13>.
- Liberman, Alvin M. – Mattingly, Ignatius G. 1985. The motor theory of speech perception revised. *Cognition* 21/1. 1–36.
- Lindblom, Björn 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35. 1773–1781.
- Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49/2. 606–608.
- Mády Katalin 2012. A fókusz prozódiai jelölése felolvasásban és spontán beszédben. In: Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 91–107.
- Mády Katalin – Bombien, Lasse – Reichel, Uwe 2008. Is Hungaraian losing the vowel quantity distiction? In Sock, Rudolph – Fuchs, Susanne – Laprie, Yves (eds.): *Proceedings of 8th International Seminar on Speech Production*. INRIA, Stras-

- bourg. 449–453. [issp2008.loria.fr/Proceedings/PDF/issp2008-106.pdf](http://issp2008.loria.fr/Proceedings/PDF/issp2008-106.pdf). (A letöltés ideje: 2018. január 5.)
- Mády, Katalin – Kleber, Felicitas 2010. Variation of pitch accent patterns in Hungarian. In: *Proceedings of Speech Prosody* 2010. <http://speechprosody2010.illinois.edu/papers/100924.pdf>. (A letöltés ideje: 2016. november 6.)
- Mády Katalin – Reichel, Uwe – Szalontai Ádám 2017. A prozódiai prominencia (nem)jelölése a németben és a magyarban. *Általános Nyelvészeti Tanulmányok* XXIX. 77–98.
- Magdics Klára 1965. *A magyar beszédhangok akusztikai szerkezete*. Nyelvtudományi Értekezések 49. Akadémiai Kiadó, Budapest.
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.
- Marasek, Krzysztof 1996. Glottal correlates of the word stress and the tense/lax opposition in German. In: *Proceedings of the ICSLP*. Vol. 96. 1573–1576.
- Markó Alexandra 2012. A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE BTK – MTA Nyelvtudományi Intézet, Budapest. 277–303.
- McCloy, Daniel R. 2016. *phonR: tools for phoneticians and phonologists*. R package version 1.0-7.
- Michaud, Alexis – Thi, Lan N. 2007. *Peakdet*. MatLab script. <https://github.com/covarep/covarep/tree/master/glottalsource/egg/peakdet>.
- Mooshammer, Christine 2010. Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *The Journal of the Acoustical Society of America* 127/2. 1047–1058.
- R Core Team 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Szalontai, Ádám – Wagner, Petra – Mády, Katalin – Windmann, Andreas 2016. Teasing apart lexical stress and sentence accent in Hungarian and German. In: Draxler, Christoph – Kleber, Felicitas (eds.): *Tagungsband 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum (P&P 12)*. Ludwig-Maximilians-Universität, München. 216–219.
- Thèoret, Hugo – Pascual-Leone, Alvaro 2002. Language acquisition: Do as you hear. *Current Biology* 12/21. pR736–R737. [http://www.cell.com/current-biology/pdf/S0960-9822\(02\)01251-4.pdf](http://www.cell.com/current-biology/pdf/S0960-9822(02)01251-4.pdf). (A letöltés ideje: 2018. január 28.)
- Tolcsvai Nagy Gábor 2017. Jelentéstan. In Tolcsvai Nagy Gábor (szerk.): *Nyelvtan*. Osiris Kiadó, Budapest. 205–499.
- Varga László 1985. Intonation in the Hungarian sentence. In Kenesei, István (ed.): *Approaches to Hungarian. Volume one. Data and descriptions*. JATE, Szeged. 205–224.
- Varga László 2000. A magyar mellékhangsúly fonológiai státusáról. *Magyar Nyelvőr* 124. 99–108.
- Varga László 2012. Van-e magyar mellékhangsúly? In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest. 35–49.

**Some articulatory and acoustic parameters of accented and unaccented vowels in Hungarian**

In the present study three members of the Hungarian vowel inventory (/i/, /u/, /ɒ/) were analysed as a function of prominence, with respect to gender and vowel quality. The theoretically most prominent (stressed and accented) and non-prominent (unstressed and unaccented) realizations were compared in terms of duration,  $f_0$ , formants, and OQ (measured with two different methods). The last two of these parameters were estimated and analyzed systematically for the first time in the study of Hungarian speech.

There was a significant interaction between the effect of prominence and vowel quality: prominence led to longer duration for the vowels /ɒ/ and /i/, but had no significant effect on /u/. We found a three-way interaction between prominence, vowel quality and gender, due to different patterns observed between the two genders in the case of the vowel /i/. Formant analysis based on Euclidean distance from the vowel space centroid did not reveal any significant effect of the degree of prominence. The comparison of  $F_1$  and  $F_2$  values showed considerable differences between the accent conditions in the case of the second formant of /ɒ/. For the two measures of OQ, we found different patterns for genders and vowels: prominence led to higher OQ values for women and lower OQ values for men. These between-gender differences were the most salient for the vowel /ɒ/.

## HIÁTUSOS KONTEXTUSOK TEMPORÁLIS MINTÁZATA GYERMEKEK BESZÉDÉBEN

Gósy Mária

MTA Nyelvtudományi Intézet

### Bevezetés

Az anyanyelv-elsajátítás folyamatában a beszédhangok képzésének fejlődése abban ragadható meg, hogy a gyermekek mind pontosabban képesek az adott fonéma realizációjára egyrészt a hangminőséget, másrészt a szegmentális és a szupraszegmentális időviszonyokat létrehozni (Hawkins 1984; Zharkova et al. 2011; Redford 2013; Gósy 2005; Bóna–Imre 2010; Auszmann 2014, 2015; Krepsz 2015; Tar 2017; Bunta et al. 2018). Mindez nem független sem az adott nyelv fonéma- és beszédhangkészletétől, sem a fonológiai folyamatok érvényesülésétől. A csecsemő már a gőgicsélés során igen gyorsan összekapcsolja a hangkezdeményeket, a felnőtt környezet pedig felismerni véli az anyanyelvi magánhangzók, illetve a magánhangzók és a mássalhangzók kapcsolatait. Ezek és a hosszabb hangsorok már valamiféle koartikulációs működést igényelnek, a szükséges beszédképző szervek közel tudatos kontrollját (Farnetani–Recasens 1999; Alcock 2006; Iverson 2010).

A holofrázisok időszakában a hangsorok jelentést kapnak, és a funkció megjelenése további összerendezést és egyre pontosabban működő kontrollt igényel a gyermektől. A felnőtt nyelv szavainak leképezése sajátosan történik, rendszerint bizonyos artikulációs egyszerűsítések révén. Ezek leggyakrabban a hasonló beszédhangok megjelenését (pl. gyermeknyelvi hasonulások), illetve beszédhangok kiesését, azaz zéró ejtését jelentik (pl. Vértes O. 1953; S. Meggyes Klára 1971).

A következő anyanyelv-elsajátítási szakaszban a beszédhangok koartikulációja fokozottan előtérbe kerül, ahogy a gyermek mentális lexikona növekszik. Ezek a jelenségek nemegyszer sajátosan valósulnak meg. Közülük most a beszédhangok betoldásaira fókuszálunk a kiejtett szavakban, amelyek a gyermek önálló – egyszeri vagy huzamosabb időn át fennálló – produkciójának eredményei. Az okok – az adott életkortól és a nyelvi szinttől függően – többfélék lehetnek. Jó részük (főként az egyszeri ejtések) feltételezhetően véletlen ejtés következményei, a felnőtt nyelv terminusával élve, artikulációs

nyelvbottlások. Más részükre az jellemző, hogy a betoldás lényegében ejtés-könnyítőként funkcionál, amikor az adott szavak morfológiaiilag komplexebbé válnak, például *papilot*<sup>4</sup> ('papírt'), *tányélot* ('tányért'), *jadijot* ('radírt'), *dógozottál* ('dolgoztál'). Ezekben az esetekben a gyermek a mássalhangzókapcsolatot oldja fel a betoldott magánhangzóval, és így teszi könnyebbé a maga számára a kiejtést. Nem zárható ki azonban az sem, hogy ezek a példák valójában a szótó és a toldalék létrehozásának szabályalkalmazását tükrözik: az *egér* ~ *egeret*, *hal* ~ *halat*, *vár* ~ *várat*, *szék* ~ *széket* típusúak hatására. A gyermeknyelvben tapasztalható betoldások egy másik része éppen nem könnyíti, hanem látszólag nehezíti az artikulációt. Ezek lehetnek perszeverációs (pl. *köpke* – 'köpeny'; *temptomp* – 'templom'; *vankond* – 'vakond') vagy anticipációs jelenségek (pl. *elengáncs* – 'elegáns'; *ungájunk* – 'ugrálunk'). Az első esetben nem törölődnek az aktivált és kiejtett beszédhangok, a másodikban pedig egy később ejtendő beszédhang korábban jelenik meg a hangsorban. Nem minden betoldás magyarázható a beszédtervezési sorrendiséggel, például *sengícscs* – 'segíts', *aupiókat* – 'autókat', *lám pánt* – 'lámpát'; *kalampács* – 'kalapács'. Ezeknek a példáknak jó része hosszabb időszakban is tapasztalható a gyermek ejtésében. Jellegzetes, hogy a leggyakrabban nazális, illetve bilabiális mássalhangzók betoldása történik, amelyek nehezíteni látszanak az artikulációt. Lehetséges, hogy csupán gyermeknyelvi egyszerű nyelvbottlásoknak tekinthetők, amelyek aztán hosszabb időszakban fennmaradnak. Összességében a betoldások az anyanyelv-elsajátítás egy szakaszára jellemző jelenségek, amelyek nem megakadásjelenségek, mint a felnőtt nyelvben (Gyarmathy 2017), hanem az artikulációs kontroll átmeneti hiányának következményei.

A gyermek kiejtésének, folyamatos beszédének és az anyanyelvi stratégiáknak a kialakulására és a fejlődésére hatással van az artikulációs kontroll és a koordináció. Lenneberg már több mint fél évszázada leírta (és azóta sokan igazolták), hogy az artikulációs mozgások működtetése önmagában nem elegendő a nyelvfejlődéshez (1967). A megfelelő artikulációs mozgások releváns tényezője az általános motoros ügyesség és annak fejlődése a nyelvelsajátítás során (Iverson 2010). Az anyanyelv elsajátításának folyamatában a gyermeknek elsősorban meg kell tanulnia koordinálni és kontrollálni az artikulációs szerveit, mivel ez alapvető feltétel a hangzó nyelv működtetéséhez. Hozzájárul ehhez az is, hogy az élet első másfél évében el kell sajátítania számos finommozgást, amelyek nélkül nem alakul ki az életkor- és nyelvspecifikus kiejtés (pl. Darrah et al. 2003). Ez azt (is) jelenti, hogy az artikuláció szükségessé teszi az ajkak, a nyelv, a szájjpadlás, a fogazat igen

---

<sup>4</sup> A példák két gyermek beszédéből valók, életkoruk a szavak ejtésekor 1;8 és 2;4 közötti volt (a szerző saját gyűjtései).



gyors, pontos és összerendezett mozgásait, összekapcsolva mindezeket a zöngképzéssel és a légzéssel. Ez a koordináció tipikus fejlődésben hároméves korra már elvárható (Iverson 2010). A koordináció és a kontroll folyamatai szoros összefüggésben működnek a felsőbb kognitív működésekkel, amelyek mintegy kölcsönhatásban eredményezik a hangzó nyelv megfelelő kialakulását. Az artikuláció, az artikulációs mozgások koordinációja és az artikulációs gesztusok kontrollja életkor-specifikusan változik, és egy-egy jelenség tekintetében szemléletesen vizsgálható. Ilyen például a két magánhangzó kapcsolatának, a hiátusnak az ejtése, amelynek kivitelezése során az artikulációs gesztusokon túl számos egyéb készség és működés szükséges ahhoz, hogy a gyermek a két szomszédos magánhangzó ejtését nyelvspecifikusan valósítsa meg.

Két magánhangzó egymás utáni előfordulásának jelenségét, azaz a köztük lévő űrt a fonetikában és a fonológiában hiátusnak nevezik akkor, ha a szomszédos magánhangzók két szótagmagot alkotnak, de nem jön létre diftongus (Siptár 2002; Britain–Fox 2009; Gósy 2014). A hiátus feloldásának módzatai nyelvspecifikusak, vagyis az egyes nyelvekre jellemző, hogy a két magánhangzó kapcsolatát miként valósítják meg, például diftongussal, epentézissel, az egyik magánhangzó kiesésével, siklóhang (glide) létrehozásával, a magánhangzók vegyülésével, illetve asszimilációval (vö. részletesebben: Gósy 2014). Ugyanazon nyelvben a hiátus feloldására különböző stratégiák is létezhetnek, mint a magyarban is. A magyar beszédben a hiátus feloldására a fonológiai megközelítésben három lehetőség kínálkozik: a magánhangzók egyikének törlése, egy hiátustöltő hang közbeékelése, avagy a két magánhangzó egyikének átalakítása félmagánhangzóvá (Siptár 2002). A hiátustöltő megjelenésének kritériumát a magyarban úgy határozzák meg, hogy az akkor következik be, ha a magánhangzó-kapcsolat egyik tagja az [i] vagy az [i:] magánhangzó. A hiátustöltés szóban és szóhatáron egyaránt bekövetkezik. A fonológiai megállapítás szerint az [e:] -vel alkotott kapcsolatokban a hiátustöltő megjelenése opcionális, rendszerint egyénfüggő (Siptár–Törkenczy 2000).

Más magánhangzók kapcsolataiban tehát nem jelenik meg hiátustöltő. Ez utóbbiakban a magánhangzó-kapcsolatok ejtése többféle lehet. Gósy (2018) kutatásában a nem [i]-t, [i:] -t, avagy [e:] -t tartalmazó magánhangzó-kapcsolatok esetén háromféle ejtési (hiátusfeloldási) stratégiát tapasztalt: (i) az egyik magánhangzó törlődik a kapcsolatból (pl. *de aztán* ejtése: [dɒsta:n], *oda elmentem* ejtése: [odɛlmɛntɛm]), (ii) a két magánhangzó összeolvad, az eredmény általában egy hosszú magánhangzó lesz (pl. a *rááll* ejtése: [ra:l:]; *mennyi időben* ejtése: [mɛn:i:dø:ben]) és (iii) átmeneti ejtés, amikor a kapcsolat első magánhangzójának ejtéséből fokozatosan alakul ki a második magánhangzó, és ez a formánsok mozgásán jól látható (pl. *kivezető ebből; például*). Ez utóbbi esetben további vizsgálatokat igényel, hogy a két magánhangzó egyike félmagánhangzóvá válik-e vagy sem (Siptár 2002).

A hiátustöltőre vonatkozóan a magyar szakirodalom többféle terminust használ, és különféle mélységekig részletezi a kérdést. Fontos tudnunk a felnőtt nyelv főbb jellemzőit a hiátustöltő ejtésére vonatkozóan, hiszen ezt a mintát hallja a gyermek az anyanyelv-elsajátítás során. A legtöbb kutató a [j]-hez hasonló mássalhangzóként jellemzi ezt a hiátustöltőt (Siptár 2002; Olaszky 2010; Menyhárt 2006; Markó 2012; Gósy 2014).

A gyermek morfológiai elemző képessége az anyanyelv-elsajátítás során sajátosan és a kezdetekben nyilvánvalóan korlátozottan működik. Fokozatosan képessé válik a szótó és a ragok elkülönítésére, illetőleg a toldalékolás tudatos használatára (jól tükrözi ezt a *malacval* /2;2/ példa, ahol a fonológiai koartikuláció megvalósítását a gyermek még nem tudta érvényesíteni). A felnőtt nyelvi fonológiai folyamatok eredménye egy ideig mintakövetéssel tükröződik a gyermek ejtésében, például jáccanak, kébzeld, zölcséglevest, felébrettem (életkor: 1;10–2;4). (A koartikuláció észlelésére vonatkozóan lásd Zamura et al. 2016.)

A szomszédos magánhangzók is előfordulnak már a korai nyelvelsajátítási szakaszoktól a gyermek beszédében. A felnőtt nyelv hiátustöltő [j]-betoldásai értelemszerűen megjelennek a gyermeknyelvi szavakban, azaz a gyermek nyilvánvalóan utánozza a hallott mintát, például *apájé, maji* ('mai'), *kamijon, Felícija, teját* (1;10 és 2;4 közötti életkorban adatoltak). A /j/ fonéma realizációi közelítőhangként viszonylag koraiak a gyermekek beszédében (Gósy 2005). Példák: *haj, baj, bújjunk, ujjá, játszik, jéggel, cipőjét, újság, vágja, olyan, autópálya, furulyál, helyére, szomjas* stb. (1;8 és 2;4 közötti időszakból). Az elemzések szerint az artikulációjuk tipikus fejlődésben már három-, négyéves korra stabilnak tekinthető (Tar 2017). Ez azért releváns a hiátustöltő tekintetében, mert a [j] artikulációs konfigurációját a gyermeknek adott kontextus(ok)ban kell kiejtenie egy más funkcióban. Kérdés, hogy vajon a funkcionális különbség együttjár-e valamilyen temporális eltéréssel is. Mind a hiátustöltőt, mind a /j/-realizációkat tartalmazó szavak a hallottak leképeződései a gyermekek beszédében. Valószínűsítjük, hogy a gyermek számára nem különböző a *fíjú* és a *bújik* szavak észlelése. A kétszavas frázisok időszakában a mintakövetés ténye a szóhatáron megvalósuló, gyakran észlelt esetekre is fennáll.

Felmerül azonban a kérdés, hogy a szavakban, az [i, i:] -t (esetleg [e:] -t) megelőzően vagy követően beékelődő [j]-szerű hangot a gyermek artikulációs könnyebbségként, mintegy ösztönösen ejti, avagy a felnőttektől hallott szóalakot ismétli, csakúgy, mint a /j/-realizációkat tartalmazó szavak esetében. A kérdés még bonyolultabb akkor, ha a szóhatáron beékelődő [j]-szerű hang ejtésére gondolunk, ami egyes esetekben persze szintén értelmezhető – mint fentebb láttuk – mintakövetésnek. Vannak olyan szókapcsolatok, amelyek – elsősorban gyakoriságuk miatt – a gyermek mentális lexikonában eleinte mint egyetlen szó tárolódnak, ilyen például a *mijaz* ('mi az'), *mijez* ('mi ez'), *hójan* ('hol van'). Ezt olyan közlések támasztják alá, mint: *anuka az*

*mijez* (1;11). A szókincs növekedése lehetővé teszi a szavak véletlenszerű összekapcsolódását a gyermek közléseiben, és gyakorlatilag lehetetlen megmondani, hogy az adott ejtés mintakövetés eredménye, avagy a gyermek önálló „alkotása”. Vajon a *víziló is* szókapcsolat (*víziló(j)is*) a gyermeki észlelés, avagy a felismert nyelvi szabály alkalmazásának a következménye? Mindez természetesen akkor merül fel, amikor a gyermek az adott szavakat folyamatosan ejti, tehát nincs semmilyen szünet a két egymást követő szó ki-ejtése között (Newton–Wells 1999).

A nyelvi fejlődés abban is jelentkezik, hogy a gyermeknek fel kell ismernie a hiátustöltés szabályát, vagyis azt, hogy mely magánhangzós kontextusokban szükséges a hiátustöltő [j] ejtése, és melyekben nem. A szóhatáron megjelenő hiátusok ejtése már egyértelműen a gyermek kreatív nyelvhasználatát tükrözheti, hiszen nem feltételezhető, hogy minden szókapcsolat mintakövetés eredménye lenne a beszédében. A gyakori szókapcsolatok észlelése, például *mijez* ('mi ez') alapján levont szabály(ok) alkalmazását valószínűsítjük a gyermek feltételezeten önálló (azaz nem utánzó) szókapcsolatainak létrehozásában, például [i]-s környezetben: *szobába(j)is*, *papa(j)ide*, *bácsi(j)ott*, *ne(j)írájál*, *víziló(j)is* (életkor: 1;8–2;0). Általában nem tapasztalható a hiátustöltés a korai anyanyelv-elsajátítás időszakában sem azokban az esetekben, ahol az nem jellemző a felnőtt nyelvre, például: *beütöttem*, *maci is*, *jóutat*, *ide állítom*, *apuka elvittem*, *megitta Andisz* ('Andris'), *olló apának*, *bácsi inekj* ('énekel'), *maci iszi* ('issza'), *vágja almádat* (1;9–2;2).

Gósy kutatásában (2014) azt a kérdést vizsgálta, hogy vajon a hiátustöltők és a /j/ fonéma realizációi között csak funkcióbeli különbség van-e, avagy a funkcióban jelentkező eltérésnek vannak-e ejtési következményei a felnőttek beszédében. Eredményei szerint különbség volt kimutatható a kétféle funkcióban ejtett [j] mássalhangzó időtartamaiban, valamint az első és a második formánsok értékeiben. Az időtartambeli eltérés azonban nagyon kicsi volt, átlagosan 10 ms-nak adódott, ennyivel voltak hosszabbak a /j/ realizációi szemben a hiátustöltőkkel. A szerző egy másik kutatásában (Gósy 2018) a hiátus feloldását vizsgálta valamennyi adatolható jelenség esetében felnőttek spontán beszédében. A [j]-féle mássalhangzó betoldása, tehát amikor a magánhangzó-kapcsolat egyik tagja a felső, illetve középső nyelvállású, illabiális magánhangzók egyike volt, adódott a leggyakoribbnak (az összes hiátust tekintve 62,7%). A más magánhangzókat tartalmazó hiátusok ejtésekor a magánhangzó törlése 5,6%-ban fordult elő, a magánhangzók összeolvadása 11%-ban jött létre, míg az átmeneti ejtésnek nevezett feloldás 20,6%-ban volt adatolható.

A hiátus feloldásának elsajátítására vonatkozóan – tudomásunk szerint – alig néhány kutatás látott napvilágot az egyes nyelvekben. Az angol beszélők gyakran használják a [ɹ] mássalhangzót két szomszédos magánhangzó ejtésekor. Tizenhárom hatéves ausztráliai angolt beszélő gyermek beszédében elemezték ennek a stratégiának az érvényesülését (Yuen et al. 2017). Arra keres-

tek választ, hogy vajon a gyermekek használják-e a [ɪ]-t a magánhangzó-kapcsolat ejtésékor, és ha igen, az vajon tudatos betoldásnak tekinthető-e. Utóbbi eldöntéséhez a harmadik formáns csökkenését vizsgálták az első magánhangzóban. A gyermekek közül nyolcan produkálták a [ɪ]-t két magánhangzó között, a többiek valamiféle glottalizációt alkalmaztak. A beékeltséget jellemző volt a harmadik formáns menetének relatív csökkenése az első magánhangzóban más kontextusokhoz képest. A szerzők ezeket az eredményeket úgy értelmezték, hogy azok a gyermekek, akik alkalmazták a betoldást, tudatosan előre terveztek, és a [ɪ]-t mint a hangsor egy szegmensét artikulálták. Egy longitudinális kutatásban, ugyancsak az angol nyelv elsajátításával kapcsolatban, azt találták, hogy a vizsgált gyermek a kezdetekben (2;4 körül) a [j]-t, illetve a glide-ot használta a magánhangzó-kapcsolatokban, 3 éves korától a /r/ realizációit, és idősebb korában a glottális zár jelent meg nagyobb arányban ugyanazon funkcióban (Newton–Wells 2002).

A hiátus jelenségének felnőtt nyelvi jellemzői nyilvánvalóan hatnak a gyermek nyelvelsajátítására és nyelvhasználatára. Kérdés az, hogy az anyanyelv-elsajátítás során miként érvényesülnek a magyarra jellemző hiátusfeloldási stratégiák, azaz a hiátus ejtésére vonatkozó fonológiai megállapítások. Itt tehát arra keresünk választ, hogy miként viselkednek azok a magánhangzó-kapcsolatok a gyermekek ejtésében, amelyek a felnőtt nyelvben szükségszerűen tartalmaznak hiátustöltőt ([i]-s környezetek), illetve választható a hiátustöltős ejtés ([eː]-s környezet), avagy nem tartalmaznak hiátustöltőt. Kérdés az is, hogy a hiátustöltőt tartalmazó, illetve nem tartalmazó kontextusok időviszonyai változnak-e az életkor előrehaladtával.

Kutatásunk célja, hogy a magánhangzó-kapcsolatok realizálódásának és ejtési időtartamainak mérésével választ adjunk a fentiekben megfogalmazott kérdésekre 4 éves és 6 éves gyermekek beszédének elemzésével. A kutatás eredményeinek bemutatását megelőzően hangsúlyozzuk, hogy a jelenséget kontextusban elemezzük, azaz a két magánhangzó kapcsolatát („V<sub>1</sub>V<sub>2</sub>”), illetőleg a hiátustöltő megjelenése esetén a „V<sub>1</sub>/hiátustöltő/V<sub>2</sub>” szekvenciákat. Ennek a megközelítésnek elsősorban a szegmentálási bizonytalanság az oka, ugyanakkor a kontextusok temporális elemzése összevethető időértékekről nyújt információkat. Összehasonlításképpen vizsgáltunk egy harmadik kontextust is, amikor a két szomszédos magánhangzó között a szó a [j] palatális közelítőhangot tartalmazza.

Négy hipotézist állítottunk fel. (i) Mindhárom vizsgált kontextus időtartamai nagymértékű hasonlóságot fognak mutatni a 4 évesek beszédében. (ii) A hiátustöltés nélküli magánhangzó-kapcsolatok és az intervokális /j/-realizációkat tartalmazó kontextusok szignifikáns eltérést mutatnak a 6 évesek beszédében. (iii) A hiátustöltő betoldása eltérő kontextus-időtartamokat fog mutatni a két korcsoportban. (iv) A mért kontextus-időtartamok jellegze-

tesen rövidebbek lesznek az idősebb gyermekek beszédében, mint a fiatalabbakéban mindhárom kontextusban.

### Módszertan

A kutatáshoz 20 magyar anyanyelvű óvodás gyermek (egynyelvűek) spontán beszédfelvételét választottuk ki véletlenszerűen a GABI adatbázisból (Bóna et al. 2014). A gyermekek egyik csoportjába az átlagosan 4 évesek kerültek (3;11–4;2), a másikba az átlagosan 6 évesek (6;1–6;4); csoportonként 5 lány és 5 fiú. A szülők által kitöltött anamnézislapok szerint tipikus anyanyelvi fejlődésű gyermekek voltak, beszédük az elvárt időben indult. Valamennyien ép hallók, ép ejtésűek, szociális háttérük hasonló, budapestiek.

Az adatbázisból azokat a felvételeket választottuk ki, amelyekben a gyermekek családjukról, kedvenc játékaikról, meseélményeikről relatíve folyamatosan beszélnek. Mintegy 100 percnyi (több mint 1,6 órnyi) spontán beszédet elemeztünk. Adatközlőnként átlagosan 5 perc volt a felvételi idő, amely azonban az interjúkészítő közléseit is tartalmazta. A gyermekek saját beszédidejének átlaga 2 perc és 50 mp volt (az adott témáról), a legrövidebb 2 percig, a leghosszabb 3 perc 20 másodpercig tartott.

A vizsgálatokhoz három kontextust határoztunk meg a gyermekek beszédében. Kettő hiátusjelenség volt, ezeket aszerint különítettük el, hogy ékelődött-e hiátustöltő a két magánhangzó közé, avagy nem. Azokban a kontextusokban, ahol a hiátustöltő közbeékelődött, a kis elemszám miatt az elemzések során nem különítettük el azt, hogy az [i], az [i:], illetve az [e:] magánhangzók V<sub>1</sub>, avagy V<sub>2</sub> helyzetben voltak-e. Míg a felső nyelvállású, palatális magánhangzók szomszédságában kivétel nélkül megjelentek a hiátustöltők, addig anyagunkban az [e:] környezetében csupán 2 esetben fordult elő hiátustöltő (mindkettő hatévesek ejtésében). Szegmentáltuk továbbá a /j/ fonéma realizációit tartalmazó kontextusokat (V<sub>1</sub>/j/V<sub>2</sub>). A kontextusok azonosítását, tehát azt, hogy a két magánhangzó között megjelenik-e a hiátustöltő, és hogy az a [j] szegmens-e, egy fonetikus és a szerző végezte. Nem volt a döntéseik között eltérés. A vizsgált anyagban a két magánhangzó között az összes [i], illetve [i:] és két esetben az [e:] magánhangzó környezetében volt tapasztalható a [j] hiátustöltő. Nem különítettük el a szavakban, illetve a szóhatáron megvalósuló kontextusokat, részben a kis elemszám miatt, részben azért sem, mivel a felnőtt nyelvben sem mutattak ki közöttük akusztikai-fonetikai különbséget (vö. Gósy 2014). Tekintettel a spontán gyermeknyelvi közlésekre, a kontextusokban megjelenő magánhangzók minősége nem volt kontrollált paraméter. A különböző magánhangzók ugyanakkor igen hasonló arányban jelentek meg a kontextusokban, a két korcsoportban. Mind a négy-, mind a hatéveseknél az [i], [i:] magánhangzók lényegesen gyakoribbak voltak a magánhangzó-kapcsolat első tagjaként, mint második tagként. Az 1. ábra az elemzett kontextusokat alkotó magánhangzók arányait szemlélteti a fiatalabb és az idősebb gyermekek közléseiben. Az [ø] és az [y] magánhangzókat nem

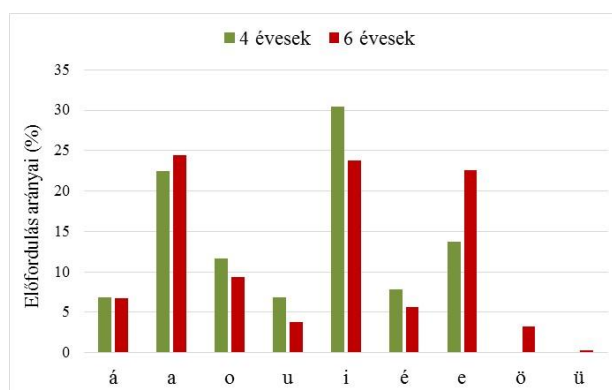
adatoltuk a négyéveseknél. A magánhangzó-minőségek előfordulása a hiátus első, illetve második magánhangzójaként hasonló arányokat mutatott mindkét korcsoportban.

A háromféle elemzett kontextus a következő volt (példákkal szemléltetve, amelyekben a hiátust és a /j/-s kontextust félkövérítettük):

$V_1+V_2$ : amikor a két magánhangzó kapcsolatába nem ékelődik egy további hang, pl. *autópálya*, *például*, *beütötte*, *videó*; *vásárolniis*, *haelkapod*, *meseaz*.

$V_1+(i)$ , ill.  $(i)+V_2$ , azaz [i]-t, [i:] -t tartalmazó (ill. [e:] -s) környezet: amikor a két szomszédos magánhangzó egyike a felsorolt három magánhangzó valamelyike, pl. *múmi<sup>j</sup>ásat*, *amij<sup>j</sup>óta*, *logopédij<sup>j</sup>ázok*, *focij<sup>j</sup>edzésem*, *barátnőj<sup>j</sup>immel*; *nyomni<sup>j</sup>és*, *menni<sup>j</sup>akkor*, *támadó<sup>j</sup>is*.

$V_1+/j/+V_2$ : amikor két magánhangzó között jelenik meg a /j/ fonéma realizációja mint a szavakat felépítő szegmentum, pl. *tojásokat*, *bújócskába*, *hülyéskedünk*, *iskolábajárni*, *egyrejebb*, *manójó*.



1. ábra

A három kontextusban előforduló valamennyi magánhangzó arányai az életkor szerint

A rögzített beszédanyagokat szó-, illetve az adott kontextus szintjén annotáltuk folyamatos akusztikai visszacsatolás és vizuális ellenőrzés mellett a Praat programban (Boersma–Weenink 2015), függetlenül attól, hogy a vizsgálni kívánt szegmentumok szóban vagy szóhatáron fordultak-e elő. A szegmentálást, illetve az annotálást egy fonetikus és a szerző végezte. Összesen 221 magánhangzó-kapcsolatot adatoltunk, 51-et a négyévesek és 170-et a hatévesek ejtésében. A három vizsgált kontextusban az előfordulások többé-kevésbé kiegyenlítettnek tekinthetők (1. táblázat). A négyévesek anyagai rövidebbek voltak, ezt a vizsgált jelenségek előfordulása is tükrözi. A szóban megjelenő kontextusok száma a 4 éveseknél 28 volt, a szóhatáron átívelőké 23; a 6 éveseknél az előbbiek száma 82, az utóbbiaké pedig 88.

A magánhangzók időtartamát a második formáns megjelenésétől annak lecsengéséig mértük a fonetikában szokásos kritériumok mentén (a méréseket a szerző és egy másik fonetikus végezte). Ha szükséges volt, figyelembe vettük az első formáns változását és a rezgéskep információját is. Az első kontextusban két magánhangzó időtartamát, a második és a harmadik kontextusban pedig két-két magánhangzó, a hiátustöltő, illetve a /j/-realizáció időtartamát, tehát összesen három-három beszédhangot mértünk. Az adatokat egy erre a célra létrehozott szkript segítségével automatikusan nyertük ki.

1. táblázat. A vizsgált beszédhang-kapcsolatok előfordulása a kontextus és az életkor szerint ( $V_1$  = első magánhangzó,  $V_2$  = második magánhangzó; 100%-nak az összes előfordulást tekintettük az adott életkorban)

Életkor	A vizsgált kontextusok előfordulása					
	$V_1+V_2$		$V_1 + [i, i:], [e:]), \text{ ill. } [i, i:], [e:] + V_2$		$V_1+/j/+V_2$	
	db	%	db	%	db	%
4 évesek	12	23,5	21	41,2	18	35,2
6 évesek	58	34,1	38	22,4	74	43,5

A három kontextus időtartamait (mint függő változókat) a következő tényezők (mint független változók) mentén elemeztük: (i) a kontextus típusa és (ii) a gyermekek életkora; elemeztük továbbá a típus és az életkor interakcióját is. A statisztikai elemzéseket a GLMM módszerrel az SPSS 20.0 szoftver alkalmazásával végeztük. A megbízhatósági szintet 95%-ra állítottuk be.

### Eredmények

Tudjuk, hogy a /j/ fonéma realizációi viszonylag korán megjelennek a gyermekek beszédében, és hogy az artikulációjuk három-, négyéves korra stabilnak tekinthető (Gósy 2005; Tar 2017). Ezért kézenfekvőnek látszott annak elemzése is, hogy az intervokális helyzetben ejtett [j] mássalhangzós kontextusok és a hiátustöltés [j]-szerű mássalhangzóját tartalmazók mutatnak-e hasonlóságot az időzítésben. Az anyanyelvi beszédészlelés (mint írtuk, két fonetikus megítélése alapján) hasonlóan, illetve azonosnak találta a jelen kutatásban vizsgált négyévesek és hatévesek ejtésében hallható [j] mássalhangzókat, funkciójuktól függetlenül.

### A kontextusok előfordulása

A  $V_1+V_2$  típusú kontextusok, azaz azok a magánhangzó-kapcsolatok, amelyek nem tartalmaztak hiátustöltőt, lényegesen gyakoribbak voltak a hatéve-

sek beszédében, mint a négyévesekében. Jellemző az is, hogy míg a fiatalabbaknál mindössze 7-féle ilyen típusú kontextust adatoltunk, addig az idősebeknél 18-félét. A fiatalabbaknál előfordulók a következők voltak (többségükben 2-2 példát találtunk rájuk a közléseikben): *au*, *ea*, *ae*, *ua*, *ii* és *ao*, egy esetben pedig az *óé*, ekkor azonban hiátustöltőt nem ejtett a gyermek. Az előfordulások mintegy fele szóhatáron történt. Amennyire megítélhető, valamennyi szó relatíve gyakori használatának feltételezhető az e korúak beszédében (pl. *autó*, *fiú a*, *de ebbe*, *de egyik*). A hatéveseknél adatolt, ugyanezen típusú magánhangzó-kapcsolatok a következők voltak: *eo*, *ea*, *ae*, *au*, *áu*, *ii*, *aa*, *ee*, *eő*, *aó*, *óa*, *aő*, *aá*, illetve öt esetben a kapcsolat második tagja az [e:] magánhangzó volt: *eé*, *aé*, *ué*, *óé*, *áé*. Az utóbbiak összesen tíz esetben fordultak elő, és mindössze kétszer adatoltunk hiátustöltőt az ejtésben: *anyu[j]ék* és *de[j]én*. Ezeket természetesen a hiátustöltős kontextusoknál elemeztük. Kutatásunkban mindkét életkori csoportban azonosak és a leggyakoribbak az *ea* és az *ae* kapcsolatok voltak, és mind a 4, mind a 6 éveseknél előfordultak továbbá az *ii*, az *aó* és az *óé* hiátusok. A szavak, amelyek ezekben a kapcsolatokban érintettek voltak, általában gyakorinak valószínűsíthetők, például: *ha elkapod*, *vele az*, *sárga akkor*, *ha ott*, *utálja azt*, *videójátékot*. Előfordultak azonban olyanok, amelyek ritkának tűntek vagy egyedinek egy-egy gyermek közlésében, például: *támadó én*, *nyerésre aztán*, *újraélesztik*, *megölte egy*. A szavak relatív gyakoriságának még az óvatos minősítése is rengeteg problémát hordoz, azonban az időzítéssel kapcsolatos eredmények lehetséges magyarázatai sokszor nem függetleníthetők a gyakoriság tényétől. Ezért tértünk ki ennek a tényszerű megemlítésére.

A fonológiai megállapítás szerinti, kötelezően hiátustöltőt tartalmazó kontextusokat a gyermekek minden esetben, életkortól függetlenül így is ejtették, amely – mint írtuk – az anyanyelvi beszédészlelés számára egy [j]-féle más-salhangzónak felelt meg. Egyetlen esetben sem ejtettek ugyanakkor hiátustöltőt, amikor a felnőtt nyelv hiátusfeloldási szabályai szerint hiátustöltés nincsen. Az opcionálisan ejthető [e:]s környezet esetében mindössze kétszer fordult elő a hiátustöltő megjelenése.

#### A kontextusok időviszonyai

Az elemzések a háromféle kontextus teljes időtartamaira koncentráltak. Először azokra a kérdésekre kerestük a választ, hogy vajon az életkor mint tényező hatással van-e a vizsgált kontextusok időtartamaira, és hogy az egyes kontextus-típusok hasonló vagy eltérő időzítési mintázatokat mutatnak-e. A statisztikai elemzések szerint szignifikáns különbséget kaptunk a vizsgált kontextus-típusok függvényében az életkorra most tekintet nélkül, az összes adat figyelembevételével ( $F(2, 270) = 6,433$ ;  $p = 0,04$ ). A vizsgált gyermekek ejtésében tehát a háromféle kontextus temporális mintázata eltérő. A négyévesek és a hatévesek kontextus-időtartamainak összehasonlító elemzése ugyancsak szignifikáns eltérést mutatott ( $F(1, 220) = 10,006$ ;  $p = 0,002$ ). Az



életkor mentén tehát jelentősen különböznek az időértékek. Az életkor és a kontextus-típusok közötti interakció is szignifikáns lett ( $p = 0,032$ ), ami arra utal, hogy az eltérő kontextusok különbözőképpen hatnak az időtartamokra az életkor mentén.

Az összes mért adatot tekintve a páronkénti összehasonlítás szerint a hiátustöltőt és a /j/-realizációkat tartalmazó kontextusok között nem volt szignifikáns eltérés ( $p = 0,064$ ); a hiátustöltő nélküli kontextus és a hiátustöltőt tartalmazó kontextus, valamint a hiátustöltő nélküli kontextus és a /j/-realizációkat tartalmazó kontextusok között a különbség azonban matematikailag igazolódott ( $p = 0,035$ , ill.  $p = 0,019$ ), vö. 2. táblázat.

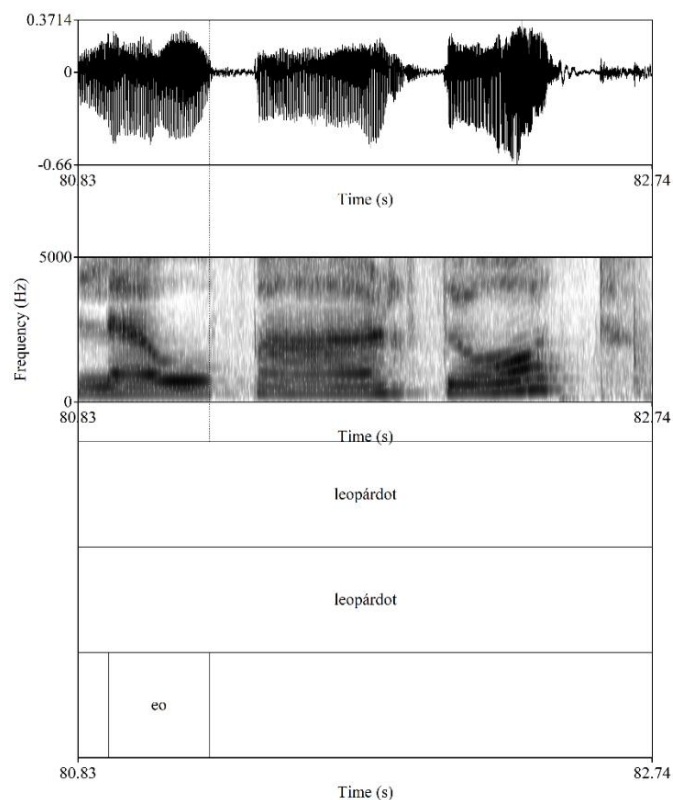
2. táblázat: A kontextusok időtartamának statisztikai alapú mátrixa

Kontextusok	Statisztikai összefüggések az egyes kontextusok között		
	V <sub>1</sub> +V <sub>2</sub>	[i]-s (ill. [e:] -s) kontextus	V <sub>1</sub> +/j/+V <sub>2</sub>
V <sub>1</sub> +V <sub>2</sub>		$p = 0,035$	$p = 0,003$
[i]-s (ill. [e:] -s) kontextus	$p = 0,035$		nincs összefüggés
V <sub>1</sub> +/j/+V <sub>2</sub>	$p = 0,003$	nincs összefüggés	

Az összes kontextus átlagos időértéke a négyéveseknél 279 ms (átlagos eltérés: 82 ms), a hatéveseknél pedig 244 ms (átlagos eltérés: 61 ms) volt, a különbség jelentősnek mondható, átlagosan 35 ms. Az egyes kontextus-típusok tekintetében csak az [i]-s (ill. az [e:] -s) kontextus esetében volt kimutatható szignifikáns különbség a 4 és a 6 éves gyermekek között ( $F(1, 59) = 18,045$ ;  $p = 0,001$ ). A másik két kontextus-típus időtartamai között nem volt matematikailag igazolható eltérés az életkor függvényében. Az életkori csoportokon belül a fiatalabb és az idősebb gyermekek kontextus-időtartamai sajátos eltéréseket mutattak. A négyévesek ejtésében a három kontextus-típus időtartamában, egymáshoz képest nem tudtunk szignifikáns különbséget igazolni ( $F(2, 50) = 4,025$ ;  $p = 0,134$ ), annak ellenére, hogy az átlagértékek jellegzetesen eltértek egymástól. (Ennek lehet oka az adatok kontextusonkénti relatíve kis száma is, illetve a nagy szóródás is.)

A fiatalabb gyermekek legrövidebben a két magánhangzóból álló kapcsolatokat ejtették, az átlagérték 250 ms (átlagos eltérés: 70 ms). Átlagosan 20 ms-mal voltak hosszabbak a /j/ mássalhangzó realizációit intervokális helyzetben tartalmazó kontextusok, az átlagérték 270 ms (átlagos eltérés: 82 ms). Leghosszabbnak a hiátustöltőt tartalmazó kontextusok adódtak, itt az átlagos időtartam 304 ms volt (átlagos eltérés: 74 ms). A három kontextust szemlélte-ti a 2., a 3. és a 4. ábra.

A hatéveseknél szignifikáns különbséget igazoltunk a kontextus típusának függvényében ( $F(2, 169) = 13,776$ ;  $p = 0,001$ ), vagyis a kontextus típusa matematikailag igazolhatóan hatással volt az időviszonyokra. A páronkénti összehasonlítás szerint a magánhangzó-kapcsolatok és a /j/-realizációkat tartalmazó kontextusok között ( $p = 0,001$ ), valamint a hiátustöltőt és a /j/-realizációkat tartalmazók között volt szignifikáns eltérés kimutatható ( $p = 0,017$ ). A magánhangzó-kapcsolatokat és a hiátustöltőt tartalmazó kontextusok időtartamai nem különböztek szignifikánsan ( $p = 0,481$ ).

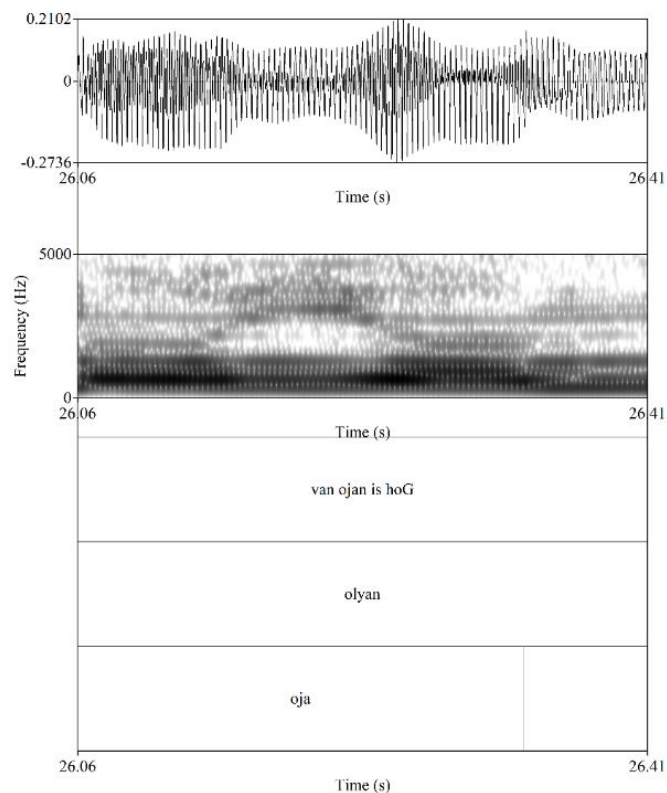


2. ábra

A magánhangzó-kapcsolatot a *leopárd* szóval szemléltető kontextus akusztikai lenyomata egy négyéves gyermek ejtése alapján

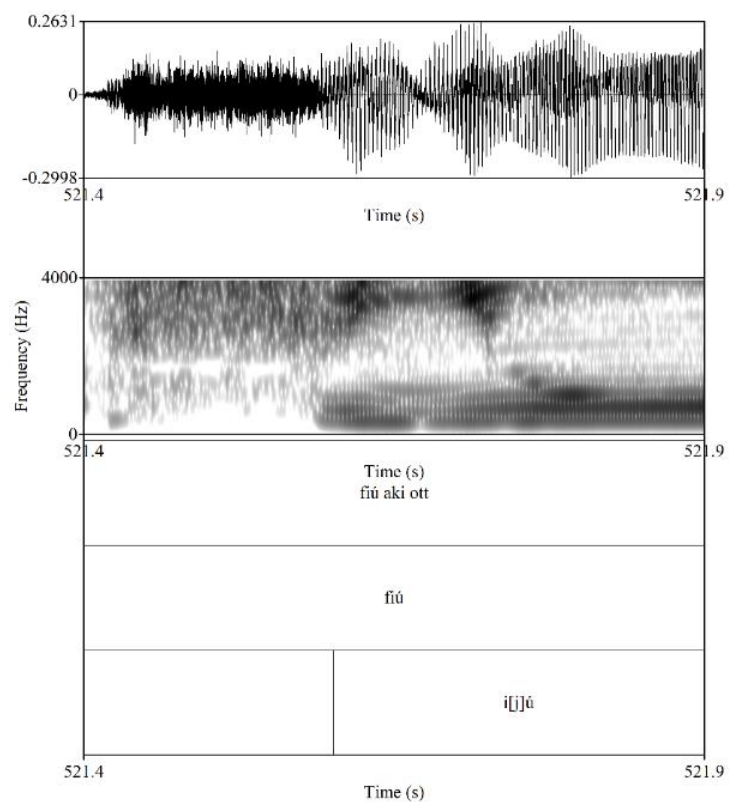
Az idősebb gyermekek legrövidebben – a négyévesekhez hasonlóan – a két magánhangzóból álló kapcsolatokat ejtették, az átlagérték ezeknél 226 ms

volt (átlagos eltérés 65 ms). A hatévesek átlagosan 263 ms-ban ejtették a /j/-realizációkat tartalmazó kontextusokat (átlagos eltérés: 59 ms), vagyis 37 ms-mal hosszabban, mint a  $V_1+V_2$  típusúakat. E tekintetben hasonló az adatok tendenciája, mint a négyévesek esetében. Az idősebbek csoportjában a hiástöltőt tartalmazó kontextusok átlagértéke 234 ms volt (átlagos eltérés 50 ms), amely mintegy 29 ms-mal rövidebb, mint a /j/-realizációkat tartalmazó kontextusok átlagértéke. Az 5., 6. és 7. ábra szemlélteti az ejtésekről készült akusztikai lenyomatokat.



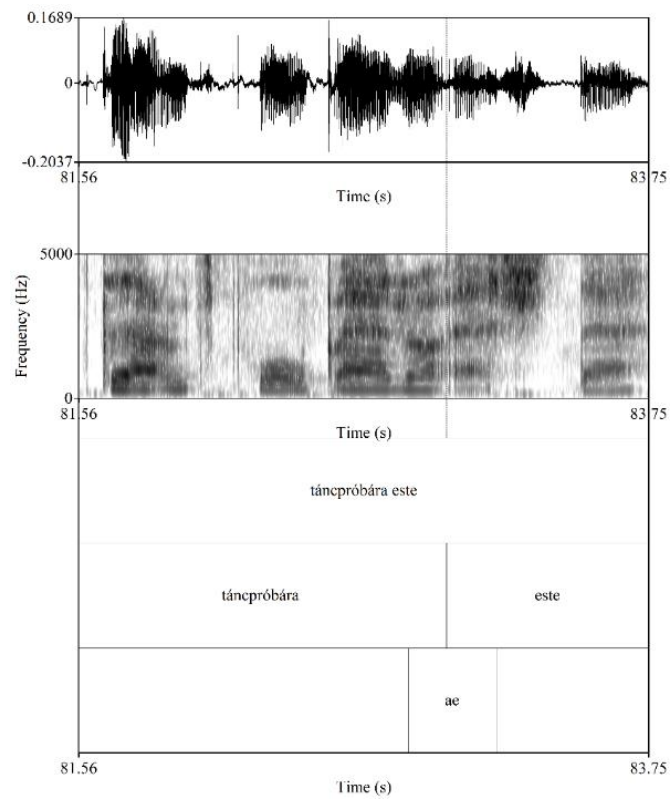
3. ábra

A [j] mássalhangzót intervokális helyzetben tartalmazó *olyan* ([ojɒn]) szó akusztikai lenyomata egy négyéves gyermek ejtésében



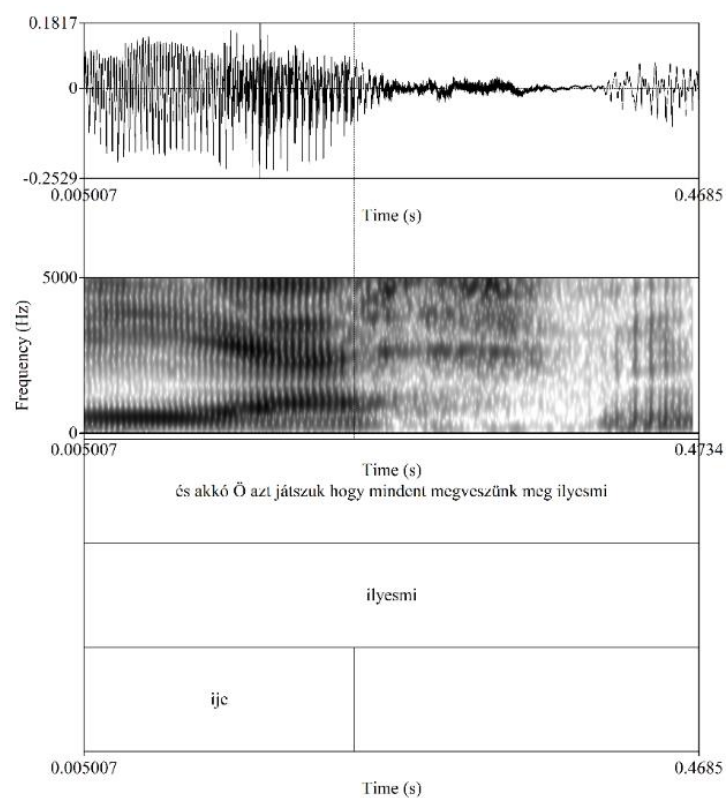
4. ábra

Hiátustöltőt tartalmazó kontextus akusztikai lenyomata a *fiú* ([fiju:]) szóban egy négyéves gyermek ejtése alapján



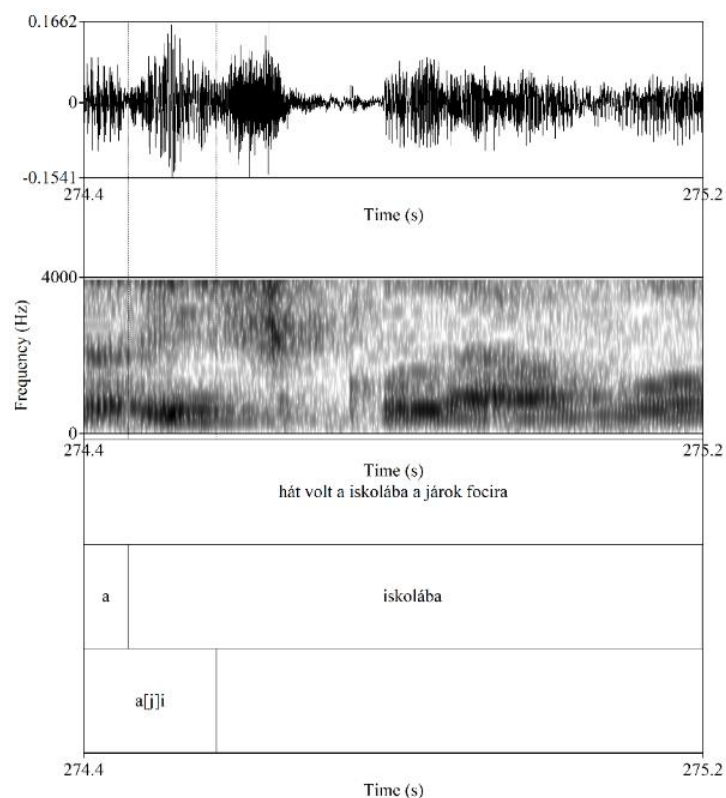
5. ábra

A magánhangzó-kapcsolatot a *táncpróbára este* példával szemléltető [ɔɛ] kontextus akusztikai lenyomata egy hatéves gyermek ejtésében



6. ábra

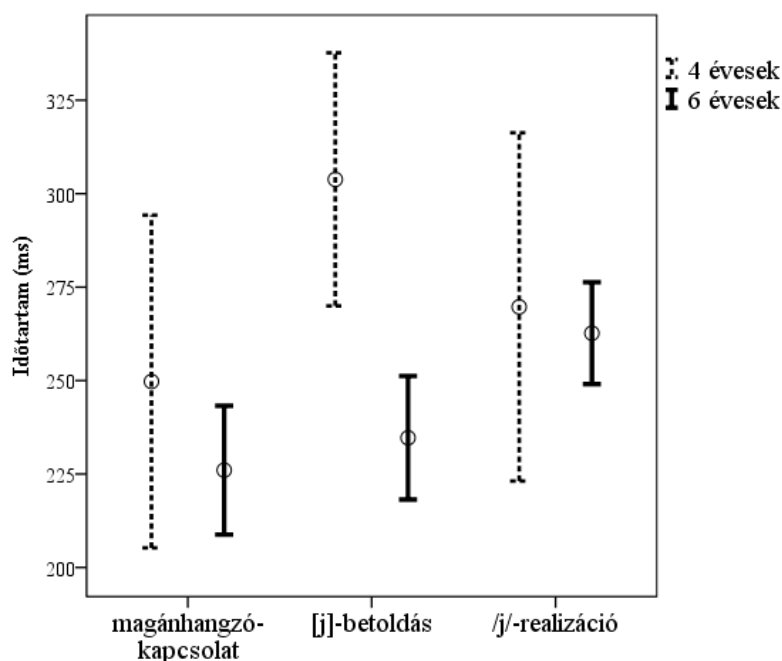
A [j] mássalhangzót intervokális helyzetben tartalmazó *ilyenmi* ([ijɛʃmi]) szó akusztikai lenyomata egy hatéves gyermek ejtésében



7. ábra

A hiátustöltőt tartalmazó *a iskolába* (ejtése: [ɔjiʃkola:bɔ]) szó rezgésképe és hangszínképe egy hatéves gyermek ejtése alapján

Azt láttuk tehát, hogy az egyes kontextus-típusok temporális viszonyai közötti különbségek sajátosan alakulnak a két életkori csoportban. A /j/-realizációt tartalmazó kontextusok átlagértékei alig térnek el az életkor függvényében, az idősebbeké mindössze 7 ms-mal rövidebb a fiatalokénál. Nagyobb a különbség a magánhangzó-kapcsolatok időtartamaiban, az átlagok szerint az idősebb gyermekek 24 ms-mal rövidebben ejtették őket. Az [i]-s és [i:] -s (ill. [e:] -s) kontextusok átlagos időtartama a négyéveseknél 70 ms-mal volt hosszabb, mint a hatéveseknél. A két korcsoport között tehát szembevető eltérés tapasztalható a hiátustöltőt tartalmazó kontextusok időtartamaiban. Az átfedések a két életkori csoport adataiban azonban mindhárom kontextus-típusban jellegzetesek. Az elemzett kontextusok időviszonyait (átlag és szó-ródás), valamint az életkorok közötti átfedéseket szemlélteti a 8. ábra.



8. ábra

A vizsgált kontextusok időtartamai négyéves és hatéves beszélők spontán közléseiben (átlag és szóródás)

Jellegzetesen különböznek az időadatok a minimum- és a maximumértékek tekintetében mindhárom kontextus esetében (3. táblázat); nagyobb tartományban realizálódnak a négyévesek, mint a hatévesek ejtésében. A fiatalabbaknál 120 ms és 457 ms között, azaz 337 ms-os tartományban, az idősebeknél pedig 101 ms és 395 ms között, azaz 294 ms-os tartományban. Valamennyi kontextus időtartamainak minimumértékei rövidebbek az idősebb gyermekek ejtésében; két kontextus-típusnál pedig a maximumértékek is náluk rövidebbek.

A  $V_1+V_2$  típus esetén az idősebb gyermekek időtartamainak értéktartománya 92-ms-mal haladja meg a fiatalabbakét. A hiátustöltőt tartalmazó kontextusok időértékeinek tartománya hasonló a két életkori csoportban, a különbség mindössze 38 ms; itt a 4 évesek tartománya nagyobb. A /j/-realizációkat tartalmazó kontextusoknál ismét jóval nagyobb az értéktartományok közötti eltérés, 81 ms, és ismét az idősebbeké a nagyobb.



3. táblázat: Az elemzett kontextusok időtartamainak minimum- és maximumértékei a két életkori csoportban

Kontextus típusa	A kontextusok időtartamadatai (ms)			
	4 évesek		6 évesek	
	minimum- érték	maximum- érték	minimum- érték	maximum- érték
$V_1+V_2$	120	322	101	395
[i]-s, [i:] -s (ill. [e:] -s) kontextus	168	439	136	369
$V_1+/j/+V_2$	316	457	164	386

### Következtetések

Kutatásunkban a hiátusjelenséget elemeztük magyar anyanyelvű gyermekek beszédében, első ízben. Az összehasonlíthatóság érdekében a hiátus teljes kontextusát elemeztük, vagyis a két szomszédos magánhangzót, illetve azokat a kontextusokat, amikor hiátustöltő ékelődött a két magánhangzó közé. Vizsgálatunkat kiterjesztettük az intervokális helyzetű palatális approximáns teljes kontextusának időtartamára is:  $V_1+[j]+V_2$ .

A eredmények szerint a négyévesek és a hatévesek elemzett kontextusainak időtartamai szignifikánsan különböztek az életkor tekintetében. A négyévesek kontextusainak időtartamai hasonlóak voltak, a hatéveseknél is csak a  $V_1+V_2$  és az intervokális [j]-t tartalmazó kontextusok különültek el szignifikánsan. Ezek az adatok az első két hipotézisünket megerősítették. Valószínűsíthető, hogy a négyévesek esetében a kontextusok hasonló temporális mintázata arra utal, hogy a gyermekek artikulációs koordinációjában még nem különülnek el sem a (feltételezetten) felnőtt nyelvi mintát követő ejtések, sem pedig az önálló szabályalkalmazások. Ez összefügghet továbbá az ejtett szavak gyakoriságával, morfológiai szerkezetével, szótagszámával is. Az a tény, hogy a magánhangzó-kapcsolatok és a /j/-realizációkat tartalmazó kontextusok matematikailag is igazolhatóan elkülönültek a hatévesek beszédében, többféle magyarázatra is utalhat. Azt gondoljuk, hogy a hiátustöltő ejtése mint artikulációs stratégia az idősebb gyermekeknél már rutinszerűbb, megszokottabb, az esetek jelentős részében vélhetően tudatos. Az ejtés motoros kontrollja már jobban működik náluk, mint a fiatalabbaknál, és ez jelenik meg az [i]- és [i:] -s kontextusok (beleértve a két [e:] -kontextusos adatot) rövidebb időtartamában a /j/-realizációkat tartalmazó kontextusokhoz viszonyítva. Az ejtett szavak gyakorisága, morfológiai összetettsége és szótagszá-

ma – elméletileg – a hatéveseknél is magyarázhatja a /j/-realizációkat tartalmazó kontextusok hosszabb voltát. Az utóbbit alátámaszthatja az a tény, hogy a – vélhetően relatíve hosszabb időtartamban ejtett rövid(ebb) szavak – jelentősen eltérő arányban fordultak elő a kontextusokban. Míg a /j/-realizációkat tartalmazók 46%-a egy és két szótagból állt, addig a hiátustöltőt tartalmazóknál ez az arány 21%, a hiátustöltőt nem tartalmazóknál pedig 19% volt. Természetesen majd kontrollált méréseknek kell igazolniuk e lehetséges magyarázat helyességét.

Ha a hiátustöltőt tartalmazó kontextusokat összevetjük a két életkorban, az látható, hogy azok jelentősen rövidebbeknek adódtak a hatévesek ejtésében szemben a négyévesekéivel. Ez megerősítette az erre vonatkozó hipotézisünket. Úgy gondoljuk, hogy az átlagosan 70 ms-nyi különbség utalhat arra, hogy a hiátustöltés megvalósítása életkor-specifikus. A fiatalabb gyermekeknél a gyakori szavakban feltételezhető a mintakövetés, illetőleg a saját rutin is, és ennek megfelelően ezekben rövidebb a hiátustöltős kontextus. Azokban az esetekben azonban, ahol kvázi-önállóan kell létrehozniuk két szomszédos magánhangzó kiejtését a hiátustöltő megfelelő beillesztésével, esetleg ritkább előfordulású szavakban, ott az artikuláció nyilván még kevésbé koordinált, több idő szükséges a hiátustöltésre. Ezt a feltételezést az is megerősíteni látszik, hogy a /j/-realizációkat tartalmazó kontextusok lényegesen kisebb (és nem szignifikáns) eltérést mutatnak a 4 és a 6 évesek ejtéseinek időtartamaiban. Adataink igazolták a hipotézisünket a hiátustöltős kontextusok életkor-specifikus temporális mintázatára vonatkozóan. Csupán illusztrációként közöljük néhány hiátustöltős kontextus időértékét a két életkort szembeállítva. A *fiúk* lexémában a hiátusos kontextus időtartama egy 4 éves gyermek ejtésében 226 ms, egy 6 évesében pedig 190 ms volt. A *múmiásat* szóban 243 ms (4 éves gyermeknél), a *cseleznie* ejtésében pedig 164 ms (6 évesnél). Egy 4 éves gyermek ejtette hiátustöltős kontextus a *mi ott* szókapcsolatban 253 ms-nak, egy 6 éves ejtésében a *mi az* pedig 160 ms-nak adódott.

Az egyes kontextus-típusok közötti időkülönbségek a négyévesek és a hatévesek között valószínűsíthetően a kevésbé rutinszerű, ritkábban előforduló, esetleg éppen az önálló alkalmazás időszakára jellemző artikulációs működéseket igazolják. Az, hogy gyakorlatilag nincs különbség az intervokális [j] kontextusának időtartamaiban a két életkori csoport között, annak tudható be, hogy az ilyen jellegű közlésrészletek már a korai gyermeknyelvben is adatolhatók, tehát már a négyévesek is bizonyos rutinnal hozzák létre a szükséges artikulációs konfigurációkat. Jelentősebb eltéréseket mértünk a magánhangzó-kapcsolatok esetében, az adatok itt arra utalhatnak, hogy ezek mindkét csoportban a ritkábban ejtett kontextusok közé tartoznak. Ezt a tényt a két életkorban jelentősen eltérő előfordulásuk, valamint az egyes típusok kevés számú megjelenései is alátámasztani látszanak. Az a tény, hogy a  $V_1+V_2$  típus esetében az idősebb gyermekek maximumértékei nagyobbak voltak, mint a fiatalabbaké, magyarázható azzal, hogy ők több mint kétszer annyi ilyen (kü-

lönböző) magánhangzó-kapcsolatot ejtettek, mint a fiatalabbak. A hiátustöltőt nem tartalmazó kontextusok időtartamai között sem igazolódott szignifikáns különbség. Úgy véljük, ez elsősorban azzal magyarázható, hogy az ilyen típusú kontextusok megjelenése a mentális lexikon kiépülésével és a hosszabb tartamú spontán közlések ejtésével kapcsolatos, ami éppen a hatéves kor tájára tehető.

Némiképpen spekuláció, de mégis valószínűsíthetjük, hogy a hatéveseknél éppen ebben a korban alakul ki a magyarra jellemző distinkció az [i], [i:] (ill. [e:])-s és a nem ezeket a magánhangzókat tartalmazó hiátusok ejtésére vonatkozóan. Nem lehet véletlen, hogy a legnagyobb időtartam-különbséget éppen a hiátustöltőt tartalmazó kontextusok esetében mértük. Az adatok alapján feltételezzük, hogy a hatéveseknél ez a hiátusejtési stratégia már gyakorlatabban működik, a négyéveseknél pedig éppen a stratégia kialakulása zajlik, és az alkalmazás létrehozása történik. A kontextus-időtartamok tehát igazolhatóan az anyanyelvi fejlődés letéteményeseinek jellemző paraméterei közé tartoznak.

Láttuk, hogy a fonetikai elemzések a felnőtt nyelvben elsősorban az időtartamban mutattak különbséget a kétféle funkcióban megjelenő [j]-féle mássalhangzók esetében. A jelen kutatásban nem a mássalhangzókat, hanem a teljes kontextust vizsgáltuk, így a /j/-realizációk és a hiátustöltő [j]-féle mássalhangzók időtartamára vonatkozó adattal nem rendelkezünk. A hiátustöltőt tartalmazó kontextusok rövidebb volta a hatévesek ejtésében ugyanakkor nem zárja ki a felnőtt nyelvre jellemző temporális mintázat jelenlétét az idősebb gyermekeknél. A hiátusok ejtésének időtartama (óvatosan ugyan, főként az artikulációs tempó eltérései miatt) összehasonlítható a felnőttek spontán beszédében mért időadatokkal (Gósy 2018 alapján). A felnőtteknél a hiátustöltőt nem tartalmazó magánhangzó-kapcsolatok átlaga 148 ms-nak adódott, vagyis 102 ms-mal rövidebbnek, mint a négyévesekéi és 78 ms-mal rövidebbnek, mint a hatévesekéi. A hiátustöltőt tartalmazó kapcsolatok ejtése átlagosan 169 ms volt a felnőtteknél, ami átlagosan 135 ms-mal rövidebb a négyévesek és 65 ms-mal a hatévesek megfelelő időértékeinél.

A hiátus nyelvspecifikus ejtése az artikulációs koordináció megfelelő szintjét és a koartikulációs jellemzők követését igényli a gyermekektől. Angol anyanyelvű gyermekek beszédének vizsgálata alapján arra a megállapításra jutottak, hogy az artikulációs mozgások időbeli koordinálása mintegy 14 éves korig nem éri el a felnőttekre jellemző tartamokat (Walsh et al. 2006; Zharkova et al. 2012). A jelen kutatás eredményei azt egyértelműen igazolják, hogy jelentős fejlődés tapasztalható a négyévesek és a hatévesek között a hiátus megvalósításában; valamint hogy a hiátustöltő illesztése igényli a leg-hosszabb koordinációs időtartamot a négyéveseknél.

## Irodalom

- Alcock, Katie 2006. The development of oral motor control and language. *Downs Syndrome Research and Practice* 11. 1–8.
- Ausmann Anita 2014. Magánhangzók akusztikai szerkezete 7 és 9 éves kisiskolások spontán beszédében. In Bátyi Szilvia – Navracsics Judit – Vigh-Szabó Melinda (szerk.): *Nyelvelsajátítási-, nyelvtanulási- és beszédkutatások*. Pszicholingvisztikai tanulmányok IV. Gondolat Kiadó – Pannon Egyetem MFTK, Budapest–Veszprém. 90–98.
- Ausmann Anita 2015. A magánhangzók akusztikai szerkezete 9 és 11 éves iskolások spontán beszédében. *Beszédkutatás* 2015. 164–176.
- Boersma, Paul – Weenink, David 2015. Praat: doing phonetics by computer. <http://www.praat.org/> (Letöltés ideje: 2015. március 5.).
- Bóna Judit – Imre Angéla 2010. A rövid-hosszú magánhangzók óvodás és kisiskolás gyermekek beszédprodukcójában. In Navracsics Judit (szerk.): *Nyelv, beszéd, írás: Pszicholingvisztikai tanulmányok I*. Tinta Könyvkiadó, Budapest. 49–56.
- Bóna Judit – Imre Angéla – Markó Alexandra – Váradi Viola – Gósy Mária 2014. GABI-Gyermeknyelvi beszédAdatBázis és Információtár. *Beszédkutatás* 2014. 246–251.
- Britain, David John – Fox, Sue 2009. The regularisation of the hiatus resolution system in British English: A contact-induced 'vernacular universal'? In Filppula, Markku – Klemola, Juhani – Paulasto, Heli (eds.): *Vernacular universals and language contacts: Evidence from varieties of English and beyond*. Routledge Studies in Germanic Linguistics, Routledge. 177–205.
- Bunta, Ferenc – Neuberger, Tilda – Bóna, Judit – Markó, Alexandra – Jordanidisz, Ágnes 2018. The production of selected phonemically short versus long Hungarian vowel pairs by 5-, 6-, and 7-year-olds. In Babatsouli, Elena Ingram, David (eds.): *Phonology in Protolanguage and Interlanguage*. Equinox Publishing, London. 139–161.
- Darrah, Johanna – Hodge, Megan – Magill-Evans, Joyce – Kumbhani, Gayatri 2003. Stability of serial assessments of motor and communication abilities in typically developing infants – implications for screening. *Early Human Development* 72. 97–110.
- Farnetani, Edda – Recasens, Daniel 1999. Coarticulation models in recent speech production theories. In Hardcastle, William J. – Hewlett, Nigel (eds.): *Coarticulation: theory, data and techniques*. Cambridge University Press, Cambridge. 31–65.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2014. A palatális közelítőhang kétféle funkcióban. *Beszédkutatás* 2014. 17–40.
- Gósy Mária 2018. A hiátus feloldása spontán magyar beszédben. [Kézirat].
- Gyarmathy Dorottya 2017. Megakadásjelenségek a magyar spontán beszédben. MTA Nyelvtudományi Intézet, Budapest.
- Hawkins, Sara 1984. On the development of motor control in speech: Evidence from studies of temporal coordination. In Lass, Norman J. (ed.): *Speech and Language: Advances in basic research and practice*. Academic Press, Inc., Orlando, San Diego, New York, London. 317–374.

- Iverson, Jana M. 2010. Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language* 37. 229–261.
- Krepsz Valéria 2015. Magánhangzó-időtartamok alakulása a hangsor hossza és az életkor függvényében. In Várad Tamás (szerk.): *IX. Alkalmazott Nyelvészeti Doktorandusz-konferencia*. Budapest. 17–29. [http://www.nytud.hu/alknyelvdok15/proceedings/Krepsz\\_Valeria.pdf](http://www.nytud.hu/alknyelvdok15/proceedings/Krepsz_Valeria.pdf) (Letöltés: 2016. november 19.)
- Lenneberg, Eric H. 1967. *Biological foundations of language*. John W. and Sons, NY.
- Markó, Alexandra 2012. Boundary marking in Hungarian V(#)V clusters with special regard to the role of irregular phonation. *The Phonetician* 105–106. 7–26.
- Menyhárt Krisztina 2006. Koartikulációs folyamatok két magánhangzó kapcsolatában. *Beszéd kutatás* 2006. 44–56.
- Newton, Caroline – Wells, Bill 1999. The development of between-word processes in the connected speech of children aged between 3 and 7 years. In Maassen, Ben – Groenen, Paul (eds.): *Pathologies of speech and language advances in clinical phonetics and linguistics*. Whurr Publishers, London. 67–75.
- Newton, Caroline – Wells, Bill 2002. Between-word junctures in early multi-word speech. *Journal of Child Language* 29. 275–299.
- Olaszy Gábor 2010. Az [i] + V és V + [i] hangkapcsolódások akusztikai elemzése a hiátustöltés magyarázatához. *Beszéd kutatás* 2010. 76–81.
- Redford, Melissa A. 2013. A comparative analysis of pausing in child and adult storytelling. *Applied Psycholinguistics* 34/3. 569–589.
- Siptár Péter 2002. Hiátus. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen. 85–98.
- Siptár, Péter – Törkenczy, Miklós 2000. *The phonology of Hungarian*. Oxford University Press, Oxford.
- S. Meggyes Klára 1971. *Egy kétéves gyermek nyelvi rendszere*. Akadémiai Kiadó, Bp.
- Tar Éva 1017. *Fonológiai fejlődés, variabilitás, beszédhanghibák*. ELTE – Eötvös Kiadó, Budapest.
- Vértes O. András 1953. *A gyermek nyelve*. Bárczi Gusztáv Gyógypedagógiai Főiskola – Felsőoktatási Jegyzetellátó Vállalat, Budapest.
- Walsh, Bridget – Smith, Anne – Weber-Fox, Christine 2006. Short-term plasticity in children's speech motor systems. *Developmental Psychobiology* 48. 660–674.
- Yuen, Ivan – Cox, Felicity – Demuth, Katherine 2017. Planning of hiatus-breaking inserted /ɹ/ in the speech of Australian English-speaking children. *Journal of Speech, Language and Hearing Research* 14. 826–835.
- Zamuner, Tania S. – Moore, Charlotte – Desmeules-Trudel, Félix 2016. Toddlers' sensitivity to within-word coarticulation during spoken word recognition: Developmental differences in lexical competition. *Journal of Experimental Child Psychology* 152. 136–148.
- Zharkova, Natalia – Hewlett, Nigel – Hardcastle, William J. 2011. Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control* 15. 118–140.
- Zharkova, Natalia – Hewlett, Nigel – Hardcastle, William J. 2012. An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *Journal of the International Phonetic Association* 42/2. 193–208.

**Köszönetnyilvánítás**

A kutatás az NKFIH-K-120234 számú pályázat támogatásával készült.

**Temporal patterns of hiatus contexts in Hungarian children's speech**

Vowel hiatus is resolved through different mechanisms such as vowel assimilation, vowel deletion, glide or consonant epenthesis, as well as morphologically conditioned consonant insertion. Hiatus resolution may have different ways within the same language as it is the case in Hungarian. Phonological departure is that (i) a [j]-like filler always occurs between two neighboring vowels when one of them is an [i], (ii) there is no hiatus filling in case both vowels are labial, back or low or any combination of these. In the case of the palatal, illabial, mid vowels, hiatus filling is optional.

Our research question is whether the realization of vowel clusters in language acquisition meets the phonological claims. We intended to describe the temporal properties of the vowel clusters in 4-year-old and 6-year old children's speech. Our main hypothesis was that there would be a change in pronunciation of the various hiatus contexts and contexts containing a palatal approximants in an intervocalic position.

## A NÉMA SZÜNETEK SAJÁTOSSÁGAI ÓVODÁSOK ÉS KISISKOLÁSOK SPONTÁN BESZÉDÉBEN

**Gyarmathy Dorottya – Horváth Viktória**

MTA Nyelvtudományi Intézet

### **Bevezetés**

A beszédkutatásban a néma szüneteket többféle módon osztályozták funkciójuk szerint. A szünetet a fonetikai szakirodalomban elsőként Sweet (1890) említette, a légzéssel kapcsolta össze és „lélegzet-csoportnak” (breath-group) nevezte az egy kilégzéssel létrehozott egységet. A korai kutatások elkülönítették egymástól a beszélő tervezési nehézségeiből adódó néma szünetet, illetve a szintaktikai szerkezet határán létrejövő junktúrát (Boomer 1965, Lounsbury 1965). Egy másik korai megkülönböztetés alapja az, hogy artikulációs okok vagy beszédtervezési probléma áll az adott szünet megjelenése mögött (Goldman-Eisler 1968). A szünetek osztályozhatóak aszerint is, hogy grammatikai vagy nem grammatikai szerepet töltenek be a beszédben. Az elkülönítés ebben az esetben azon alapszik, hogy tartalmas vagy funkciószó előzi meg, illetve követi őket (Gee–Grosjean 1983). Azok a szünetek, amelyek tartalmas szó és funkciószó között fordulnak elő, általában grammatikai funkciójúak, szintaktikai vagy prozódiai határt jelölnek. A funkciószót követő és a tartalmas szót megelőző szünetek ezzel szemben egy szintaktikai/prozódiai egységen belül realizálódnak, nem-grammatikai típusúak. A spontán beszédben és a társalgásban többféle szünetet különböztet meg a szakirodalom. A pause (magyarul ’szünet’) egy beszédfordulón belüli jelkimaradás, a gap (magyarul ’rés, hézag’) a társalgási egységek közötti szünet, lehetőséget kínál a beszélőváltásokra; a lapse szintén (’kihagyás, megszünés’) jelezheti a társalgás végét (Sacks et al. 1974, Levelt 1989). A társalgásban előfordul továbbá gondolkodási vagy hatásszünet; a beszélő kiemelhet vele új információt, de diskurzusszervezői szereppel is bírhat (Esposito et al. 2007).

A beszédsszünetek funkcióinak elkülönítése attól (is) függ, hogy a kutatók mely paradigmarendszer alapján vizsgálják azokat. Bruneau (1973) kommunikációs szempontból három típusú csendet definiált: pszichológiai, interaktív és szociokulturális. A pszichológiai típusú általában nagyon rövid időtartamban, hezitációs jelenséggel vagy tempólassulással valósul meg, és azt a célt szolgálja, hogy időtartama alatt a hallgató feldolgozhassa az elhangzottakat. Az interaktív szünet ennél általában hosszabb időtartamú, a személyek közötti kapcsolatban játszik szerepet, például a beszélőváltások lebonyolítására szolgál. A szociokulturális szünetek egyesítik az első két típus sajátossá-

gait. Zellner (1994) a szünetek kétféle osztályozási rendszerét különböztette meg: 1. a fizikai és nyelvészeti osztályozást, és 2. a pszichológiai és pszicholingvisztikai osztályozást. Az első csoportosítás szerint a beszéd-szünet lehet intra-szegmentális vagy inter-lexikális, míg a második kategóriarendszer néma és kitöltött szüneteket is megkülönböztet. Pragmatikai szempontú elemzésében Kurzon (2007) négyféle csendet különböztetett meg: társalgási, tematikus csendet (egy témával kapcsolatban a beszélő nem hajlandó beszélni, például politikai jellegű interjúban), de más típus az is, amikor a társalgási helyzetben egy vagy több résztvevő magában elolvas valamit – például egy osztálytermi helyzetben a tanár instrukciójára a diákok elolvasnak/átfutnak egy részt a tankönyvből. A negyedik típusú a szituációs csend, például egy koncert hallgatása vagy közös megemlékezés közben. Zellner (1994) különféle aspektusokból definiálja a néma szüneteket: beszédtechnológiai szempontból a szünet egy olyan amplitúdóval nem rendelkező egység, amely fizikai jelenség; lehet egy beszédhang része (például zöngétlen zárhangok néma fázisa) vagy megjelenhet szavak között. Pszicholingvisztikai szempontból a néma szünetek együtt járhatnak kilégzéssel, nyeléssel, hangos belégzéssel.

A kutatások szerint a beszédhelyzet, a szünet funkciója, gyakorisága és időtartama összefüggést mutat. Minél komplexebb egy beszédfeladat, minél nagyobb kognitív erőfeszítést igényel, annál gyakoribb és hosszabb szüneteket tartanak a beszélők (Goldman-Eisler 1968, Kowal et al. 1975). Politikai beszédben a néma szünetek gyakrabban és hosszabb időtartamban valósultak meg, a leghosszabb szünetek stilisztikai funkciót tölthettek be – erre a beszéd-típusra nem voltak jellemzőek a kitöltött szünetek, amelyek interjúhelyzetben kimondottan gyakoriak voltak (Duez 1982). Összefüggést találtak továbbá angol nyelvű beszélőknél a szünet pozíciója és időtartama között például a 'to+infinitive' nyelvtani szerkezetek esetében. A felolvasásokban a *to* elemet megelőző szünetek szignifikánsan hosszabbak voltak, mint az azt követők; míg a spontán beszédben ennek ellenkezője igazolódott feltehetően a beszédtervezés sajátosságai miatt (Bada–Genç 2008).

A szünetek realizációinak és funkcióinak elemzésével óvodás és kisiskolás gyermekek spontán beszédében kevesebb kutatás foglalkozott a felnőttekéhez képest. A nemzetközi pszicholingvisztikai szempontú elemzések főként a gyermek életkora és a beszéd-típus függvényében elemezték a szünetek realizációját. A néma szünetek időtartama szignifikáns csökkenést mutatott 4 és 8 éves kor között (Singh et al. 2007). Ezek hosszát azonban nemcsak a gyermek életkora befolyásolja, hanem a beszéd típusa, a feladathelyzet jellege is. Óvodás és kisiskolás gyermekek szignifikánsan nagyobb százalékban tartottak szünetet kép alapú történetmesélés során annál, mint amikor társalgásban vettek részt (Deputy et al. 1982). Egy másik kutatásban szintén kisiskolás gyermekek szignifikánsan hosszabb szüneteket tartottak akkor, amikor egy



történetet szó szerint kellett visszamondaniuk annál, mikor csak a történet lényegét kellett összefoglalniuk (Schönplflug 2008).

Pragmatikai szempontból a gyermekek beszédében előforduló néma szüneteket elemezték 15 órányi anyagban, olasz kisiskolások körében. A szüneteknek az osztálytermi kommunikációból eredő funkcióját elemezték, mint például a 'várakozási idő' – a tanár kérdésére a tanuló válaszolni készül és a gondolkodáshoz időre van szüksége (Maroni 2011). Angol anyanyelvű általános iskolások osztálytermi kommunikációját vizsgálva az adatok azt mutatják, hogy a szünetezés és a beszélőváltások hasonlóan működnek, mint a többi beszédhelyzetben, de a 'várakozási időből' adódó hallgatások hosszabbak (Ingram–Elliott 2014).

A magyar beszédkutatásban is nagy hagyománya van a szünetek elemzésének. Balassa József már a 19. sz. végén tárgyalta a jelenséget, elsősorban fiziológiai szempontból (1886). Hegedűs Lajos (1953) szintén hangsúlyozta a szünet szerepét a közlésekben, felfogása szerint azonban a légzést nem a fiziológiai, hanem a gondolkodási folyamatok vezérlik a beszéd során. A magyar fonetikai szakirodalomban sokféle meghatározás és osztályozás született a jelenségre a huszadik században, ezek tartalmaztak produkciós és perцепciós szempontokat is (Fónagy 1967; Szende 1979; Váradi 1988). Sallai és Szende (1995) spontán szövegekben előforduló szüneteket tárgyalt, meghatározásuk magában foglalta a néma és a kitöltött szünetet, a zero időtartamút, illetőleg az ún. „szünetkompenzáció” eseteit (pl. néma szünet előtti magánhangzónyúlás). A szünet az általuk képviselt tágabb elméleti keretben a szekvenciának, azaz a szeriális szerkezetnek valamilyen információt képező, vagy hordozó megszakításaként értelmezhető. A beszédben ugyanakkor olyan jelkimaradások is előfordulnak, amelyek egyes beszédhangok képzéséhez köthetők (pl.: zöngétlen explozívák és zöngétlen affrikáták zárszakaszai), s mint ilyenek, nem tekinthetők beszédsszünetnek (vö. Gósy 2004).

A nagyobb mennyiségű spontán beszédet tartalmazó felvételek és adatbázisok az elmúlt évtizedben lehetővé tették a néma szünet sok szempontú vizsgálatát magyar nyelven is. A kutatások megállapították, hogy a jelenség gyakorisága és időtartama függ a beszélő személyétől (életkor, nem), a beszédhelyzettől, a témától, a beszédtypustól (vö. pl. Gósy 2000; Gocsál 2001; Menyhárt 2003; Imre 2005; Markó 2005; Olasz 2005; Laczkó 2009; Váradi 2010; Bóna 2013; Neuberger 2014).

A spontán beszéd vizsgálata azt mutatta, hogy néma szünetek gyakran jelennek meg megakadásjelenségek környezetében, különösen akkor, ha a beszélő változtat vagy javít a közlésben. A szünetek azonban nem befolyásolják a megakadások észlelését, inkább a megakadás típusától függ, hogy a hallgatónak feltűnik-e a közlésben fellépő diszharmonia (Bóna 2006). A szón belül tartott szünetek a beszédtervezés különféle problémáját jelzik; leggyakrabban toldalékmorféma előtt szakad félbe a produkció – ez a grammatikai tervezés, illetve a lexikális előhívás nehézségére utal. Az újraindításokban adatolt né-

ma szünetek időtartama hosszabb volt a szón belül tartott jelkimaradásokhoz képest (Gósy 2010, 2012). A beszédszakaszok első felében előforduló szón belüli szünetek hosszabbak voltak, mint a szakasz második felében adatoltak (Gósy–Krepsz 2017). Nagy mennyiségű spontán beszéden készült kutatás néma és kitöltött szünetek összefüggésével kapcsolatban azt mutatta, hogy a legnagyobb arányban a hezitálást követően adatolható néma szünet a beszédben. Legkevesebb volt azon esetek aránya, amikor a beszélők a kitöltött szünet előtt és után is tartottak néma szünetet. A hezitálást megelőző néma szünetek hosszabbak voltak, mint a kitöltött szünetet követőek (Horváth 2014).

Gósy (2000) a beszédszünetek funkcióinak beszédprodukciónak szempontú elemzése mellett magyar nyelven az elsők között hangsúlyozta a szünetek feldolgozásban betöltött szerepét. Az eredmények azt mutatták, hogy a hallgatók a szünetek mintegy kétharmadát észlelik; a szünet időtartama és észlelése szoros összefüggést mutat.

A magyar anyanyelvű óvodás és iskolás gyermekek beszédszüneteit az elmúlt években kezdték nagyobb adatmennyiségen vizsgálni. 6–13 éves gyermekek spontán beszédét vizsgálva Neuberger (2014) azt találta, hogy az életkor nem befolyásolta szignifikáns mértékben a néma szünetek percenkénti előfordulását. Az óvodásoknál és a kilenc éveseknél átlagosan 22,5 darab; a hétéveseknél 19,8 db, a tizenegy éveseknél 22,9 db, a tizenhárom éveseknél pedig 21,4 db néma szünet realizálódott percenként. A szünetek aránya átlagosan 30–35% volt az egyes életkori csoportokban, az adatok egyénenként 15% és 46% között szóródtak. A néma szünetek időtartama az egyes életkori csoportok között szignifikánsan különbözött: a legrövidebb néma szünetek a tizenhárom évesekre, a leghosszabbak pedig a hétévesek beszédére voltak jellemzőek; illetve a lányok rövidebb szüneteket tartottak, mint a fiúk.

5, 7 és 9 éves gyermekek beszédének temporális jellemzőinek vizsgálata azt mutatta, hogy – életkortól függetlenül – a gyermekeknél nagyon hasonlóan alakult a beszédszakaszok (71–77%) és a szünetek aránya (23–29%). A szünettartás gyakorisága egyénenként óriási szóródást mutatott. Az ötévesek csoportjában 11,1–28,9 darab/perc; a hétéveseknél 5,1–27,2 db/perc; a kilenc éveseknél pedig 8,9–36,2 db/perc voltak a szélsőértékek a gyakoriság tekintetében (Vakula–Krepsz 2017).

Diszlexiás kisiskolások beszédében a kontrollcsoportéhoz képest ugyanolyan arányban adatoltak néma szünetek (az összes beszédidő 36%-a), de a nyelvi zavarral küzdő tanulók hosszabb szüneteket tartottak, mint a tipikus beszédfejlődésű társaik (Vakula 2012).

A megakadásjelenségek és a néma szünetek összefüggéseit magyar anyanyelvű gyermekek beszédében eddig kevés kutatás vizsgálta. Mészáros (2012) iskoláskorú gyermekek és felnőttek társalgásaiban elemezte a néma szüneteket a megakadásjelenségek függvényében. Alapvetően két fő szünet-típust különített el a társalgás szerkezetével összhangban: 1. a fordulón belüli, és 2. a beszélőváltáshoz köthető néma szüneteket. A fordulón belüli néma

szüneteket aszerint különböztette meg, hogy azok tagoló szerepűek-e, illetőleg megakadásjelenségekhez köthetők. Az utóbbiba az általa felállított rendszerben nem csupán a szigorúan szerkesztési szakaszként realizálódó néma szünetek tartoznak, de minden olyan jelkimaradás, amely valamiféle megakadásjelenség előtt vagy után jelentkezik. A tagoló szerepű néma szünetek közé pedig kizárólag azokat sorolja, amelyek kötőszavak előtt vagy után, illetőleg tagmondatháron fordulnak elő.

Néma és kitöltött szünetek összefüggését vizsgálták óvodások spontán beszédében (Horváth 2014). A gyermekek – a felnőttekhez hasonlóan – leggyakrabban a kitöltött szünetet követően tartottak néma szünetet; csak nagyon ritkán fordult elő, hogy a kitöltött szünet környezetében nem volt néma szünet. A kitöltött szünetet megelőző néma szünetek átlagosan rövidebb időtartamban valósultak meg.

A jelen kutatás célja óvodás és kisiskolás gyermekek spontán narratíváiban megjelenő néma szünetek osztályozása és a szünettípusok időtartamainak elemzése. Hipotéziseink szerint a) az egyes szünettípusok gyakorisága és időtartama eltérő mintázatot mutat; b) a néma szünet hosszát meghatározza a közlésben elfoglalt helye; illetve c) az óvodások és a kisiskolások szünettartási mintázatai között különbségek adatolhatók.

### **Kísérleti személyek, anyag, módszer**

A kutatáshoz 6, 7, 8 és 9 éves gyermekekkel készültek narratívák (korosztályonként 10 fő, 5 lány és 5 fiú). A 6 évesek óvodások voltak, a 7 évesek első osztályosok, a 8 évesek második osztályosok, a 9 évesek harmadik osztályosok. A gyermekek budapestiek, egynyelvűek, tipikus beszédfejlődésűek voltak, és egyiküknek sem volt hallásproblémája és beszédhibája. A felvételi protokoll és az interjút készítő személy minden esetben ugyanaz volt: a kísérletvezető arra kérte a gyermekeket, hogy meséljenek a családjukról, vagy mit szoktak csinálni az óvodában/iskolában, mit szeretnek játszani stb. A továbbiakban az interjút készítő személy csak akkor szólalt meg, amikor a gyermek elakadt és újabb kérdésre volt szükség a folytatáshoz. A felvételek a ugyanazon megszokott óvodai/iskolai környezetben készültek Sony ICD-SX700 típusú hangfelvevővel.

A kutatás anyaga 40 narratíva, a teljes korpusz időtartama 112 perc. A felvételeket a Praat 5.3 programban annotáltuk beszédszakasz és szó szinten (Boersma–Weenink 2013). A néma szüneteket a Praat szoftverben manuálisan nyertük ki folyamatos auditív és vizuális ellenőrzés mellett, a megelőző lexéma utolsó hangjának lecsengésétől a követő lexéma első hangjának kezdetéig tartó jelkimaradást címkéztük.

A korpuszban adatolt 2596 db néma szünet időtartama összesen mintegy 35 perc. Egy gyermek átlagosan 2,8 percet beszélt, amely idő alatt 65 néma szünetet tartott. Az 1. táblázatból jól látható, hogy az egyes korosztályok adatai csak minimális eltéréseket mutatnak, egyedül a 9 éves gyermekek beszéd-

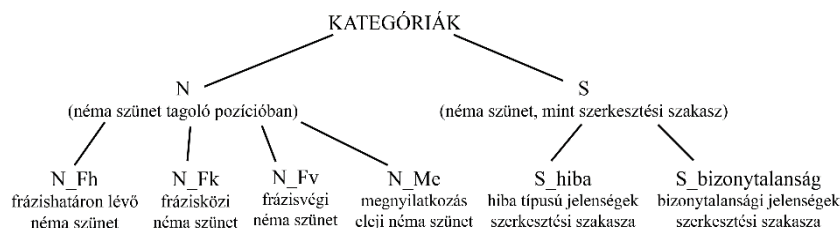
ideje mutat lényegesebb növekedést, esetükben azonban a szünetek teljes időtartama is hosszabb.

1. táblázat: A beszédidő és a néma szünetek időtartamadatai életkorok és nemek szerint

életkor	nem	teljes beszédidő (perc)	átlagos beszédidő (perc)	teljes szünetidő (perc)	átlagos szünetidő (perc)
6 évesek	fiú	11,5	2,3	3,1	0,62
	lány	12,7	2,5	3,5	0,70
7 évesek	fiú	11,8	2,4	3,7	0,74
	lány	12,3	2,5	2,7	0,54
8 évesek	fiú	13,2	2,6	3,8	0,76
	lány	12,8	2,6	5,4	1,08
9 évesek	fiú	17,2	3,4	6,8	1,36
	lány	20,8	4,2	6,1	1,22
ÖSSZES	fiú	53,7	2,7	17,4	0,87
	lány	58,6	2,9	17,7	0,89

A szünetek típusokba sorolásához Gyarmathy (2017) felnőtt beszélők narratíváihoz kidolgozott kategóriarendszerét vettük alapul. A szüneteket elsőként aszerint különítettük el, hogy azok megakadásjelenségekhez köthetők (ilyenkor az artikuláció leállításától a javítás kezdetéig tartó idő, a szerkesztési szakasz részei) vagy a tagolást szolgálják (1. ábra). A szerkesztési szakaszokban megjelenő néma szüneteket S, a tagolást elősegítő szüneteket pedig N betűvel jelöltük. Mindkét fő csoporton belül elkülöníthetők voltak további alcsoportok. A szerkesztési szakaszokat (S) aszerint kategorizáltuk tovább, hogy milyen megakadásjelenséghez tartoztak. A jelen kutatásban a következőket tudtuk elkülöníteni: a) ismétlések (S<sub>ism</sub>; *akivel nagyon S<sub>ism</sub> akivel nagyon jó*), b) újraindítások (S<sub>új</sub>; *já- S<sub>új</sub> járok rajzsakkörre*), c) téves kezdések (S<sub>tévkezd</sub>; *találtam há- S<sub>tévkezd</sub> négy német barátnőt*), d) téves szótalálások (S<sub>tévszó</sub>; *egy emeletes ágyban (házban) lakunk*, és f) szünet a szóban (SzSz; *meg az u SzSz grálós játék*). Az egyes alcsoportok nagyon kis elemszámúak voltak, ezért az adatok feldolgozása során két nagy csoportba soroltuk őket aszerint, hogy bizonytalansági, avagy hiba típusú megakadásokhoz köthetők. A tagolási pozícióban megjelenő néma szüneteket (N) pozíciójuk szerint különítettük el egymástól. Megkülönböztettük a megnyilatkozás eleji (N<sub>Me</sub>) néma szüneteket, amikor beszélőváltáskor az aktuális beszé-

lő belekezd a közlésbe; ilyenkor a szünetet legfeljebb egy töltelékszó, vagy diskurzusjelölő előzi meg: *Felvételvezető: mesélj a családdról. Adatközlő: Hát N\_Me két testvérem van.* A frázishatáron lévő (N\_Fh) néma szünetek a virtuális mondatok egyes tagmondatainak határán, gyakran kötőszó előtt vagy után helyezkednek el: *nyáron voltunk Erdélybe N\_Fh és akkor ott aludtunk.* Frázisközi (N\_Fk) szünetként jelöltük azokat, amelyek grammatikai egységen („tagmondaton”) belül fordultak elő: *az egyik nagyon N\_Fk szörnyes játék.* Frázisvégi (N\_Fv) szünetként azonosítottuk a virtuális mondatokat lezáró néma szüneteket, amely után a beszélő új virtuális mondatot kezd, gyakran egy új gondolati egységgel folytatja a közlését: *és ott majd kapok egy új kutyát N\_Fv és járok úszásra...* A frázisvégi és a frázishatáron lévő szünetek megkülönböztetése a spontán beszédben nem minden esetben egyértelmű, ezért csak az olyan virtuális mondatokat lezáró néma szüneteket azonosítottuk frázisvégiként, amelyek esetében a követő mondat nem kötőszóval kezdődik, és/vagy teljesen új gondolati egységet vezet be. Azokat az eseteket, ahol a besorolás nem volt egyértelmű, nem vontuk be az elemzésünkbe.



1. ábra

A kutatásban alkalmazott kategóriarendszer

A statisztikai elemzéseket az SPSS 20-as verziójával végeztük. Az adatainkra általános lineáris kevert modellt építettünk (GLMM), ahol a független változóink az egyes szünettípusok, a nem és az életkor, függő változóink az időtartamok voltak, random faktorként a beszélőket vettük fel. Az adatok eloszlásvizsgálatát binominális nemparaméteres, illetve Chi-négyzet goodness-of-fit tesztel végeztük, az előzetes normalitásvizsgálatot pedig a Kolmogorov-Smirnov tesztel. A szünetek időtartam-adatai nem normál eloszlást mutattak, ezért azokat logaritmikus skálára transzformáltuk és az így kapott normál eloszlású adatokon végeztük a statisztikai modell építését.

### Eredmények

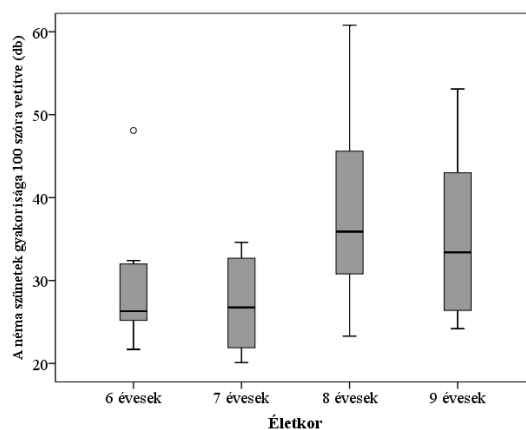
A teljes korpuszban 2596 darab néma szünetet adatoltunk. A 6 évesek beszédében összesen 527 darab, a 7 éveseknél 534 darab, a 8 éveseknél 588 darab, a 9 éveseknél pedig 947 darab néma szünet fordult elő. A szünetek per-

cenkénti előfordulása nagyon hasonlóan alakult mindhárom korosztályban (2. táblázat). A statisztikai elemzés szerint a gyermekek életkora és neme nem volt szignifikáns hatással a néma szünetek percenkénti gyakoriságára.

2. táblázat: A néma szünetek előfordulása a korpuszban

	Összes szünet száma (db)		Szünetek száma percenként (db)		Szünetek száma 100 szóra vetítve (db)	
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
<b>6 évesek</b>	53	28–88	21,7	16–29	29,1	22–48
<b>7 évesek</b>	53	30–85	21,8	14–34	27,2	20–35
<b>8 évesek</b>	59	20–131	21,6	11–29	38,3	23–61
<b>9 évesek</b>	95	49–137	25	14–33	36,1	24–53

A szünetek gyakoriságát 100 szóra vetítve is megadtuk. A 6 és 7 éveseknél; illetve a 8 és 9 éveseknél nagyon hasonlóan alakult a szünetek darabszáma 100 szóra vetítve (2. táblázat és 2. ábra). Az idősebb gyermekek gyakrabban tartottak néma szünetet a közléseikben; de a statisztikai elemzés nem igazolt szignifikáns különbséget az egyes csoportok között.



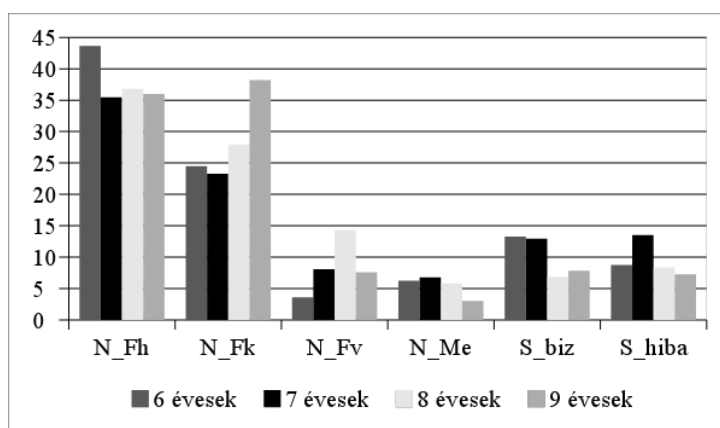
2. ábra

A néma szünetek gyakorisága 100 szóra vetítve (db)

Elsőként a két fő kategóriánkat, a tagoló pozícióban megjelenő (N) és a szerkesztési szakaszként (S) realizálódó néma szüneteket elemeztük. A két

csoport egymáshoz viszonyított aránya a következőképp alakult: az elemzett néma szünetek 81,5%-a az N-nel, 18,5%-a az S-sel jelölt csoportba tartozott; a binominális nemparaméteres teszt igazolta, hogy a főkategóriák eloszlása nem véletlenszerű ( $p < 0,001$ ). A teljes beszédidőnek a tagoló helyzetű néma szünetek a 27,1%-át (18,8 db/perc), míg a szerkesztési szakaszok mindössze a 4,3%-át (4,3 db/perc) adták. A négy életkori csoportban elemezve az adatokat, a következőképp alakulnak az arányok: A 6 évesek néma szüneteinek 78%-a volt tagoló helyzetű (17 db/perc), 22%-a szerkesztési szakasz (4,8 db/perc), az előbbiek a teljes beszédidő 22,7%-át, az utóbbiak 5%-át tették ki. A 7 éveseknél hasonlót tapasztaltunk; a tagoló szünetek aránya 73,6% (16,3 db/perc), a szerkesztési szakaszoké 26,4% (5,9 db/perc); az előbbiek a teljes beszédidő 20%-át, az utóbbiak 6,4%-át adják. A 8 és 9 éveseknél eltölteni látszik az arány, eredményeik jobban hasonlítanak a felnőtteknél adatoltakhoz. A 8 éveseknél a tagoló pozíciójú néma szünetek és a szerkesztési szakaszok aránya 84,9% (19,2 db/perc), illetve 15,1% (3,4 db/perc), míg a 9 éveseknél 85,7% (21,4 db/perc) és 14,3% (3,6 db/perc). A 8 éveseknél az előbbiek a teljes beszédidő 31,1%-át, az utóbbiak 4,5%-át tették ki; a 9 éveseknél a 31,8%-át, illetve 2,4%-át. A binominális nemparaméteres teszt mind a négy csoportban igazolta, hogy a főkategóriák eloszlása nem véletlenszerű ( $p < 0,001$ ).

Elemeztük a szünetek különböző típusainak eloszlását (3. ábra). A 6 évesek beszédében a néma szünetek az esetek döntő többségében frázishatáron jelentek meg. A frázisközi szünetek aránya ehhez képest már csak fele. A szerkesztési szakaszként realizálódott néma szünetek gyakrabban érintettek bizonytalansági megakadásokat, mint hibákat. A szerkesztési szakaszként realizálódott néma szünetek gyakrabban jelentek meg bizonytalanságokhoz, mint hibákhoz kapcsolódva. A típusok eloszlása szignifikáns mértékben különbözik a véletlenszerű eloszlástól ( $\chi^2$  goodness of fit teszt:  $\chi^2(5) = 361,125$ ;  $p < 0,001$ ). A 7 évesek korpuszában nagyon hasonló arányok adatolhatók a típusok előfordulásának tekintetében. A típusok eloszlása ebben a csoportban is szignifikáns mértékben különbözött a véletlenszerű eloszlástól ( $\chi^2$  goodness of fit teszt:  $\chi^2(5) = 191,461$ ;  $p < 0,001$ ). A 8 éveseknél szintén a frázishatároló szünetek aránya volt a legmagasabb; ők azonban nagyobb arányban tartottak szünetet frázis végen, mint a két kisebb csoport tagjai. A szünet típusok eloszlása a 8 éveseknél sem véletlenszerű ( $\chi^2$  goodness of fit teszt:  $\chi^2(5) = 291,571$ ;  $p < 0,001$ ). A 9 éveseknél a legnagyobb arányban frázisközi szünetet adatoltunk, a típusok eloszlása ebben a csoportban is szignifikánsan különbözött a véletlentől ( $\chi^2$  goodness of fit teszt:  $\chi^2(5) = 652,141$ ;  $p < 0,001$ ).



3. ábra

A néma szünetek aránya funkciójuk szerint az egyes korcsoportokban

A néma szünetek időtartama óriási szóródást mutatott. Volt olyan, amelyik 50 ms alatt realizálódott, illetve olyan is előfordult, hogy a gyermekek több mint 10 s hosszú szünetet tartottak például frázis végen, vagy a kísérletvezető kérdését követően megnyilatkozás elején, a válaszon és annak nyelvi formáján gondolkodva. A 6 évesek korpuszában 45 ms volt a legrövidebb, és 9525 ms a leghosszabb néma szünet időtartama. A 7 éveseknél a szünetek időtartama 35 ms és 11633 ms között szóródott. A 8 éveseknél 69 ms és 8870 ms voltak a szélsőértékek; a 9 éveseknél pedig 31 ms és 14513 ms volt a szünetek legrövidebb és leghosszabb időtartama.

Elemeztük a szünetidőtartamok szélsőértékének megvalósulását a szünetek típusa szerint is (3. táblázat). A legtöbb extrém hosszú néma szünet – a gyermekek életkorától függetlenül – frázisvégi és megnyilatkozás eleji pozícióban fordult elő. A szünetek időtartamának alsó határa is a frázisvégi és megnyilatkozás eleji helyzetben volt a legmagasabb. A legkevesebb extrém hosszú időtartamú szünetet a szerkesztési szakaszként realizálódó néma szünetek esetén adatoltuk.

3. táblázat: A szünetidőtartamok szélsőértékei (ms) a funkció szerint az egyes életkori csoportokban

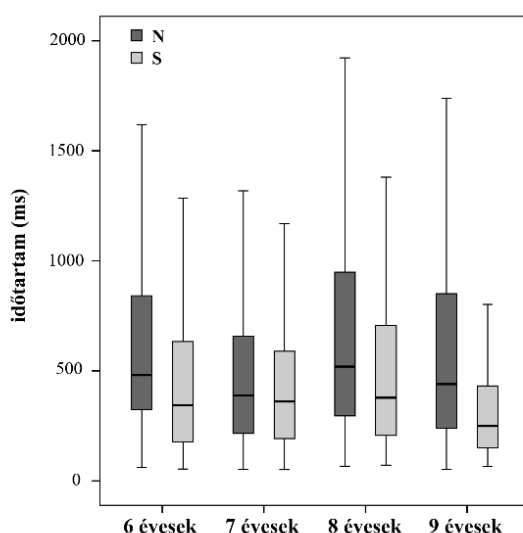
		6 évesek	7 évesek	8 évesek	9 évesek
N_Fh	Minimum	45	52	66	52
	Maximum	4223	7853	5424	10998
N_Fk	Minimum	61	42	69	58



<b>N_Fv</b>	<b>Maximum</b>	3179	4565	8870	4034
	<b>Minimum</b>	194	128	135	239
<b>N_Me</b>	<b>Maximum</b>	9525	4407	6129	11795
	<b>Minimum</b>	144	154	121	179
<b>S_biz</b>	<b>Maximum</b>	6372	11633	4953	14513
	<b>Minimum</b>	54	35	78	31
<b>S_hiba</b>	<b>Maximum</b>	3403	6270	2082	2120
	<b>Minimum</b>	65	51	71	65
	<b>Maximum</b>	1481	5757	6760	706

A szünetidőtartamok nem csak a típusok függvényében szóródtak nagymértékben, az individuális különbségek is óriásiak voltak. A hatévesek csoportjában például akadt olyan gyermek, aki a megnyilatkozás elején az esetek 90%-ban 2 s-nál is hosszabb néma szüneteket tartott; de olyan is, akinél csak egyetlen egy alkalommal adatoltunk szünetet megnyilatkozás elején.

A statisztikai elemzésnél nem vettük figyelembe a kiugró értékeket (az esetek 9,2%-a), illetve azokat az eseteket, mikor a gyermek beszédében az adott típusú szünetre csak egy példa fordult elő. Az adatok előzetes normalitásvizsgálata megerősítette, hogy azok nem normál eloszlásúak, így a statisztikai elemzések előtt a mért értékeket logaritmikus skálára transzformáltuk. A két fő kategória átlagos időtartamát elemezve elmondható, hogy a néma szünetek időtartamát szignifikánsan meghatározza az, hogy szerkesztési szakaszként, vagy tagoló szünetként fordul-e elő, amit a statisztikai elemzések is megerősítettek:  $F(1, 214) = 37,864$ ;  $p < 0,001$ ). A szünetek hosszúságát befolyásolja még ezen kívül az életkor és a nem ( $F(3, 214) = 5,106$ ;  $p = 0,002$ ) illetve a szünettípus és a nem ( $F(3, 214) = 4,807$ ;  $p = 0,003$ ) együttes hatása. A két fő típus közül minden csoportban a tagoló pozíciójú néma szünetek realizálódtak hosszabb időtartammal (4. ábra). A 6 éveseknél ez átlagosan 663 ms-ot (átlagos eltérés: 562 ms), a 7 éveseknél 529 ms-ot (átlagos eltérés: 454 ms), a 8 éveseknél 689 ms-ot (átlagos eltérés: 530ms), míg a 9 éveseknél 638 ms-ot (átlagos eltérés: 593 ms) jelentett. A szerkesztési szakaszok a 6 éveseknél átlagosan 541 ms (átlagos eltérés: 602 ms), a 7 éveseknél 477 ms (átlagos eltérés: 414 ms), a 8 éveseknél 501 ms (átlagos eltérés: 407 ms), a 9 éveseknél pedig 386 ms (átlagos eltérés: 398 ms) hosszúak voltak. A statisztikai elemzés csak a két felsőbb korcsoportban igazolta a fő szünetkategóriák szignifikáns különbségét (8 évesek:  $F(1, 214) = 10,616$ ;  $p = 0,001$ ; 9 évesek:  $F(1, 214) = 37,762$ ;  $p < 0,001$ ).



4. ábra

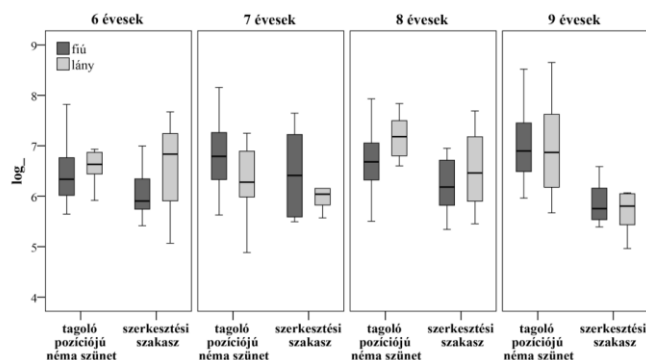
A tagoló pozíciójú (N) és a szerkesztési szakaszként (S) realizálódó néma szünetek időtartamértékei korcsoportonként

Az életkor önmagában nem volt hatással a kétféle szünet időtartamára, tehát a tagoló szünetek, és a szerkesztési szakaszok időtartamának egymáshoz viszonyított aránya konstansnak mondható, az életkor előrehaladtával nem változik; minden életkorban a tagoló pozíciójú néma szünetek a hosszabbak. A két típus időtartamait külön-külön vizsgálva az egyes korcsoportoknál azonban már elmondható, hogy mind a tagoló pozíciójú néma szünetek ( $F(3, 214) = 4,027$ ;  $p = 0,008$ ), mind a szerkesztési szakaszként realizálódók ( $F(3, 214) = 2,873$ ;  $p = 0,037$ ) hosszúságát befolyásolja a beszélő életkora. A szerkesztési szakaszok hossza csökkenő tendenciát mutat az életkor előrehaladtával. A páronkénti összehasonlítás szerint a tagoló helyzetű néma szünetek időtartama a 6 és a 8 éveseknél ( $t = 2,043$ ;  $p = 0,042$ ), a 6 és a 9 éveseknél ( $t = 2,042$ ;  $p = 0,042$ ), illetve a 7 és a 8 éveseknél ( $t = 2,798$ ;  $p = 0,006$ ), és a 7 és a 9 éveseknél ( $t = 2,756$ ;  $p = 0,006$ ) matematikailag igazolhatóan különbözik. A 6 évesekhez képest (663 ms) a 8 évesek hosszabb (689 ms), a 9 évesek rövidebb (638 ms) szüneteket tartottak; a 7 évesekénél (529 ms) pedig mind a 8, mind a 9 évesek tagoló helyzetű néma szünetei hosszabbak voltak. A leg-hosszabb tagoló szünetek a négy korosztály közül a 8 éveseket jellemezték.

A szerkesztési szakaszok időtartama pedig a 6 és a 9 évesek ( $t = 2,306$ ;  $p = 0,022$ ), a 7 és a 9 évesek ( $t = 2,273$ ;  $p = 0,024$ ), illetve a 8 és a 9 évesek között ( $t = 2,542$ ;  $p = 0,012$ ) mutatott szignifikáns különbséget. A 9 évesek

eredményei tehát mind a három további korcsoporttól különböztek, náluk adatoltuk ugyanis átlagosan a legrövidebb szerkesztési szakaszokat (386 ms). A diszharmonia feloldása a legtöbb időt (541 ms) a 6 éveseknél vett igénybe, őket követték a 8 évesek 501 ms-mal, és a 7 évesek 477 ms-mal.

Az adatokat az egyes korcsoportokon belül nemenkénti bontásban elemezve (vö. 5. ábra) a 6 ( $F(1, 214) = 5,114$ ;  $p = 0,025$ ), a 7 ( $F(1, 214) = 5,710$ ;  $p = 0,018$ ) és a 8 éveseknél ( $F(1, 214) = 5,546$ ;  $p = 0,019$ ) mutatható ki szignifikáns különbség a fiúk és a lányok között. Az összes elemzett néma szünet időtartamát tekintve a 6 (fiú: 619 ms, lány: 654 ms) és a 8 éveseknél (fiú: 579 ms, lány: 765 ms) a lányokat jellemezte a hosszabb szünettartás; a 7 éveseknél (fiú: 612 ms, lány: 438 ms) a fiúkat. A 9 évesek (fiú: 615 ms, lány: 585 ms) esetében a különbség elenyésző. A tagoló pozíciójú szünetek a 6 (fiú: 699 ms, lány: 622 ms) és a 7 évesek (fiú: 630 ms, lány: 440 ms) csoportjában a fiúknál realizálódtak hosszabban, míg a 8 évesek esetében a lányoknál (fiú: 598 ms, lány: 806 ms). A 9 éveseknél alig volt eltérés a két nem között (fiú: 653 ms, lány: 623 ms). A szerkesztési szakaszok a 6 (fiú: 385 ms, lány: 785 ms) és a 8 éveseknél (fiú: 459 ms, lány: 548 ms) a lányoknál voltak hosszabban, míg a 7 (fiú: 553 ms, lány: 431 ms) és a 9 éveseknél (fiú: 440 ms, lány: 292 ms) a fiúknál. A páronkénti összehasonlítás továbbá a 7 ( $t = 2,376$ ;  $p = 0,018$ ) és a 8 éveseknél ( $t = 2,311$ ;  $p = 0,022$ ) igazolta a tagoló helyzetű néma szünet időtartamának szignifikáns eltérését a két nemnél.



5. ábra

A tagoló pozíciójú és a szerkesztési szakaszként realizálódó néma szünetek a két nemnél az egyes korcsoportokban

Az időtartamértékek elemzését az egyes altípusokra kiterjesztve elmondható, hogy a leghosszabb átlagos időtartammal a frázisvégi (1219 ms, átlagos eltérés: 734 ms) és a megnyilatkozás eleji néma szünetek (967 ms, átlagos el-

térés: 992 ms) realizálódtak; és ez volt a két legritkábban előforduló típus is. A frázishatáron tartott néma szünetek 657 ms (átlagos eltérés: 489 ms), a grammatikai struktúrát megtörő frázisközi szünetek 423 ms (átlagos eltérés: 314 ms), a bizonytalansági megakadások szerkesztési szakaszaként 605 ms (átlagos eltérés: 575 ms), míg a hibák szerkesztési szakaszaként megjelenők mindössze 327 ms (átlagos eltérés: 233 ms) hosszúak voltak átlagosan. A négy korcsoportban a szünetidőtartamok ugyan eltérően alakultak, de tendenciaszerű egyezések megfigyelhetők. Mindegyik korcsoportban a megnyilatkozás eleji és a frázisvégi szünetek realizálódtak a leghosszabb időtartammal; a 6 és a 7 évesek felvételeiben a megnyilatkozások elején, míg a 8 és a 9 éveseknél frázisvégen tudtuk a leghosszabb szüneteket adatolni (4. táblázat).

4. táblázat: Az egyes szünettípusok időtartamértékei korcsoportok szerint

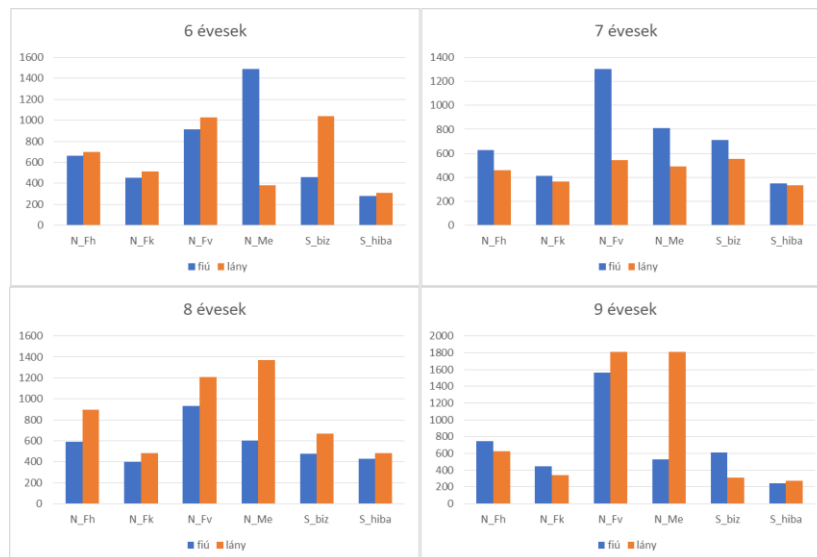
	6 évesek		7 évesek		8 évesek		9 évesek	
	Átlag (ms)	Átl. eltérés (ms)	Átlag (ms)	Átl. eltérés (ms)	Átlag (ms)	Átl. eltérés (ms)	Átlag (ms)	Átl. eltérés (ms)
<b>N_Fh</b>	677	469	520	386	722	538	680	508
<b>N_Fk</b>	489	345	390	290	442	443	403	325
<b>N_Fv</b>	985	471	948	736	1027	552	1660	791
<b>N_Me</b>	1225	1300	693	653	916	798	1111	1143
<b>S_biz</b>	702	706	623	505	544	504	520	518
<b>S_hiba</b>	290	225	338	231	463	297	258	150

Az adatok statisztikai vizsgálata kimutatta, hogy a néma szünetek időtartamát elsősorban a szünet típusa határozza meg ( $F(5, 182) = 23,555$ ;  $p < 0,001$ ), de nem független az életkortól sem ( $F(3, 182) = 2,883$ ;  $p = 0,032$ ). A szünetidőtartamot továbbá befolyásolja az életkor és a nem ( $F(3, 182) = 6,708$ ;  $p < 0,001$ ), illetve az életkor és a szünettípus ( $F(15, 182) = 2,539$ ;  $p = 0,002$ ) együttes hatása. Az egyes típusok időtartam-realizációit a vizsgált életkorok függvényében elemezve a hibák szerkesztési szakaszainál ( $F(3, 182) = 7,589$ ;  $p \leq 0,001$ ), a frázisvégi ( $F(3, 182) = 3,360$ ;  $p = 0,020$ ) és a frázisközi szünetek ( $F(3, 182) = 4,762$ ;  $p = 0,003$ ) esetében igazolódott szignifikáns különbség; ennél a három típusnál az értékek tehát matematikailag igazolhatóan is különböznek korcsoportonként (vö. 4. táblázat). A páronkénti összehasonlítás a bizonytalanságok szerkesztési szakaszainál és a frázishatáron megjelenő szüneteknél egy-egy esetben, a hibáknál két esetben, a frázisközi és a frázisvégi szüneteknél pedig három-három esetben igazolt szignifikáns különbséget (5. táblázat).

5. táblázat: A szünnettípusok páronkénti összehasonlításának statisztikai eredményei az életkorok függvényében

típus	páronkénti összehasonlítás	<i>t</i> -érték	<i>p</i> -érték
N_Fh	6 évesek – 8 évesek	2,165	0,032
N_Fk	6 évesek – 8 évesek	2,452	0,015
	7 évesek – 8 évesek	2,427	0,016
	8 évesek – 9 évesek	2,979	0,003
N_Fv	6 évesek – 9 évesek	2,147	0,033
	7 évesek – 8 évesek	2,175	0,031
	7 évesek – 9 évesek	2,897	0,004
S_bizonytalanság	6 évesek – 9 évesek	2,449	0,015
S_hiba	7 évesek – 9 évesek	2,491	0,014
	8 évesek – 9 évesek	4,232	0,000

Noha a beszélők neme önmagában nem befolyásolta a szünetidőtartamok alakulását, az életkor, a nem és a szünnettípus együttes hatása szignifikánsnak bizonyult:  $F(38, 182) = 2,191$ ;  $p < 0,001$ . A 6 évesek csoportjában a páronkénti összehasonlítás a frázishatáron megjelenő néma szünetek esetében igazolt szignifikáns különbséget a fiúk és a lányok között ( $t = 2,016$ ;  $p = 0,045$ ), a 8 éveseknél a fázishatáron megjelenő ( $t = 2,484$ ;  $p = 0,014$ ) és a frázisközi szüneteknél ( $t = 2,581$ ;  $p = 0,011$ ), a 9 éveseknél pedig a frázisközieknel ( $t = 2,194$ ;  $p = 0,030$ ). A két nem adatainak összevetése alapján általánosságban elmondható, hogy a lányok a legtöbb esetben hosszabb szüneteket tartottak (6. ábra), egyedül a 7 évesek csoportjában fordult az elő, hogy a fiúk néma szünetei voltak a hosszabbak.



6. ábra

A szünetek átlagos időtartama a két nemnél korcsoportok szerint

### Következtetések

A jelen kutatásban a beszédsszünetek realizációit elemeztük óvodás és kisiskolás gyermekek spontán narratíváiban. A fő kérdés az volt, hogy a gyermekek életkora és neme mellett a szünet pozíciója hogyan befolyásolja a temporális sajátosságokat.

A gyermekek összes közlésének 31%-a volt néma szünet. Neuberger (2014) hasonló szünetarányt adatolt óvodás és iskolás gyermekek beszédében (30–35%). A felnőtt beszélők narratíváiban kisebb volt a néma szünetek beszédhez viszonyított aránya, csak 20% (vö. pl. Gyarmathy 2017). A különbség feltehetően abból adódik, hogy a gyermekek számára – az eltérő kognitív képességek, a kevesebb beszédtapasztalat és az interjúhelyzet sajátosságai miatt – még nehezebb feladat a gondolatok nyelvi formájának megtervezése, a tervezés és kivitelezés összehangolása.

A jelen kutatásban nem volt jelentős különbség a néma szünetek percenkénti előfordulásában az egyes életkori csoportok között. Ez az eredmény hasonló egy korábbi kutatásban adatolt tendenciához, amely szerint szintén nem volt szignifikáns különbség a gyermekek között a szünetek gyakoriságában az életkor függvényében (Neuberger 2014). Az egyéni különbségek ugyanakkor óriásiak, ahogy ezt korábbi kutatások is kimutatták hasonló életkori csoportokban (Neuberger 2014, Vakula–Krepsz 2017). A 100 szóra vetített

szünet-gyakoriság két csoportra bontotta a korosztályokat: a 6 és 7 évesek, valamint a 8 és 9 évesek narratíváiban alakultak nagyon hasonlóan az adatok. Az idősebb gyermekek 100 szóra vetítve gyakrabban tartottak néma szünetet. A grammatikai komplexitás fejlődésével (Horváth 2017) több összetett virtuális mondatot hoznak létre a gyermekek, ez eredményezi egyrészt a szintaktikai eredetű szünetek gyakoriságának növekedését. A komplexebb közlések másrészt bonyolultabb beszédtervezést igényelnek; megnő a diszharmonikus jelenségek előfordulásának lehetősége is – a felszínen ez gyakran néma szünetként realizálódik.

A néma szünetek pozíciójának elemzése során elkülönítettük egymástól a tagoló helyzetű néma szüneteket (N) és a megakadásjelenségekhez kapcsolódóan megjelenőket (S), majd mindkét csoportot további alcsoportokra bontottuk. A tagoló néma szünetek aránya a gyermekek teljes korpuszában 81,5%; a szünetek 18,5%-a pedig valamilyen megakadásjelenséghez kapcsolódott. A 6 éveseknél 78%, a 7 éveseknél 73,6% volt a tagoló néma szünetek aránya. A 8 éveseknél a tagoló pozíciójú néma szünetek aránya 84,9%, míg a 9 éveseknél 85,7%; a felnőtteknél pedig 87,8% (vö. Gyarmathy 2017). A 8 és 9 éveseknél adatolt tagoló szünetek előfordulása tehát jobban hasonlít a felnőtteknél tapasztalt tendenciához, de minden vizsgált életkori csoportra igazoltuk, hogy jóval kevesebbszer tapasztalható néma szünet megakadásjelenségekhez kapcsolódva, mint tagoló pozícióban.

A tagoló pozíciójú (N) csoporton belül négy kategória szerepelt: a megnyilatkozás eleji (N\_Me), a frázishatáron lévő (N\_Fh), a frázisközi (N\_Fk) és a frázisvégi néma szünetek (N\_Fv). A 6–8 évesek narratíváiban a frázishatáron lévő néma szünetek voltak a leggyakoribbak csakúgy, mint a felnőtteknél (Gyarmathy 2017); a 9 éveseknél a frázisközi szünetre adatoltuk a legtöbb előfordulást. A gyermekek és a felnőttek (Gyarmathy 2017) beszédére is igazolható, hogy a néma szünetek nagyobb arányban jelennek meg grammatikai-  
lag is indokolt helyen, vagyis a közlés értelmi, értelmezési egységét nem megtörve (N\_Fh, N\_Fv és N\_Me), mint frázison belül. Ez arra utal, hogy a beszédtervezés során nem csupán a közlés tartalmi és formai részét tervezzük meg, de a szünettartást is (vö. Zellner 1994; Ramanarayanan et al. 2009). A grammatikai szerkezetet megtörő frázisközi szünetek hátterében valamiféle nagyobb tervezési zavar valószínűsíthető.

Az adatok azt mutatták, hogy a megakadásokhoz kapcsolódó szünetek gyakrabban köthetőek a beszélő bizonytalanságából adódó jelenségekhez, mint a hibajelenségekhez. Ez abból adódik, hogy maga a bizonytalansági megakadás – a beszélő életkorától függetlenül – jóval nagyobb arányban jelenik meg a beszédben, mint a hiba típusú jelenség (vö. pl. Gósy 2003; Szabó 2008; Bóna 2010; Neuberger 2014).

A szünetek elemzése igazolta hipotézisünket, amely szerint időtartamuk szignifikáns mértékben függ attól, hogy tagoló pozícióban vagy megakadásjelenséghez kapcsolódva jelennek meg a közlésben. A gyermekek

beszédében – életkortól függetlenül – hosszabbak voltak a tagoló jellegű néma szünetek.

A tagoló jellegű szüneteken belül mindegyik korcsoportban a megnyilatkozás eleji és a frázisvégi szünetek realizálódtak a leghosszabb időtartammal. A felnőttek narratíváiban is megnyilatkozás elején adatolták a leghosszabb szüneteket (Gyarmathy 2017). A megnyilatkozás elején ugyanis a gyermek és felnőtt beszélőknek egyaránt több időre van szüksége a gondolatok szelektálására és azok nyelvi formájának kialakítására. A frázisvégi szünetek hosszabb időtartamát pedig az indokolja, hogy az adott gondolatmenet lezárásával a közlés további részének tervezési folyamatát az elejéről kell kezdeni. Minden vizsgált életkori csoportban – a felnőttekhez hasonlóan (Gyarmathy 2017) – a frázisközi néma szünetek valósultak meg átlagosan a legrövidebb időtartamban, ezt az érthetőségre, feldolgozhatóságra irányuló kommunikációs kényszer magyarázza.

A jelen vizsgálati csoportokban a szerkesztési szakaszként realizálódó néma szünetek hossza csökkenő tendenciát mutatott az életkor előrehaladtával. Hipotézisünknek megfelelően, a különbség szignifikáns volt a hatéves óvodások és a kilencéves kisiskolások narratíváiban mért időtartamok között. Ez arra utal, hogy a nyelvi fejlődés során a gyermekek a hibajavítási stratégiákat is elsajátítják fokozatosan, így egyre rövidebb időt vesz igénybe a felmerült diszharmonia feloldása. A kognitív fejlődés mellett az iskolában a gyermekek egyre nagyobb gyakorlatra tesznek szert különféle narratívák létrehozásában (pl. szóbeli felelet, kiselőadás tartása, olvasottak összefoglalása), a nagyobb beszédtapasztalat pedig szintén hatással van a közléseikre. A szerkesztési szakaszként realizálódott néma szünetekről általános érvénnyel megállapítható, hogy a beszélő bizonytalanságára utaló jelenségek esetén sokkal több időre van szükségük a gyermekeknek a háttérben lévő diszharmonia feloldásához, mint a felszínen is megjelenő hibák esetén. A felnőtteknél is a hibák szerkesztési szakasza rövidebb volt, mint a bizonytalanságoké (Gyarmathy 2017).

A néma szünetek pozíciójának és realizációjának részletes vizsgálata a gyermekek beszédében azt mutatta, hogy már az óvodások és a kisiskolás gyermekek is hasonló szünettartási stratégiákat alkalmaznak a narratívák során, mint a felnőttek. A szünetek időtartama csökkenő tendenciát mutatott ez életkor előrehaladtával. A nyelv grammatikai struktúrája feltehetően nagymértékben meghatározza a tagolás és a szünettartás mikéntjét is. Ezt a gyermekek az anyanyelv-elsajátítás során tanulják meg, majd ahogy egyre gyakorlottabbá válnak a beszédben, a mintázatok egyre jobban idomulnak a felnőtt nyelvihez.



## Irodalom

- Bada, Erdoğan – Genç, Bilal 2008. Pausing preceding and following to in to-infinitives: A study with implications to reading and speaking skills in ELT. *Journal of Pragmatics* 40. 1939–1949.
- Balassa József 1886. *A fonetika elemei, különös tekintettel a magyar nyelvre*. Magyar Tudományos Akadémia, Budapest.
- Boersma, Paul – Weenink, David 2013. Praat: doing phonetics by computer. <http://www.fon.hum.uva.nl/praat/download.win.html>
- Boomer, Donald S. 1965. Hesitation and grammatical encoding. *Language and Speech* 8. 148–158.
- Bóna Judit 2006. A megakadásjelenségek akusztikai és perцепciós sajátosságai. *Beszéd kutatás 2006*. 101–113.
- Bóna Judit 2010. Beszédtervezési folyamatok az életkor és a beszéd típus függvényében. *Magyar Nyelvőr* 134. 332–341.
- Bóna Judit 2013. A beszéd szünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszéd kutatás 2013*. 60–76.
- Brunenau, Thomas J. 1973. Communicative silences: forms and functions. *Journal of Communication* 23. 17–46.
- Deputy, Paul N. – Nakasone, Hirotaka – Tosi, Oscar 1982. Analysis of pauses occurring in the speech of children with consistent misarticulations. *Journal of Communication Disorders* 15. 43–54.
- Duez, Danielle 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25. 11–25.
- Esposito, Anna – Stejskal, Vojtěch – Smékal, Zdeněk – Bourbakis, Nikolaos 2007. *The significance of empty speech pauses: Cognitive and algorithmic issues. Advances in Brain, Vision, and Artificial Intelligence*. Springer, Berlin Heidelberg. 542–554.
- Fónagy Iván 1967. Áthajlás, szünet, szerkezet. *Nyelvtudományi Közlemények* 69. 313–343.
- Gee, James Paul – Grosjean François 1983. Performance structures: a psycholinguistics and linguistics appraisal. *Cognitive Psychology* 15. 411–458.
- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszéd kutatás 2001*. 61–72.
- Goldman-Eisler, Frida 1968. *Psycholinguistics: Experiments in spontaneous speech*. Academic Press, London.
- Gósy Mária 2000. A beszéd szünetek kettős funkciója. *Beszéd kutatás 2000*. 1–15.
- Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2010. Az önellenőrzés egy jelensége a spontán beszédben. *Acta Hungarica* 19-20. 194–207.
- Gósy Mária 2012. Az artikuláció leállása a spontán beszédben. In Navracsics Judit – Szabó Dániel (szerk.): *Mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III*. Tinta Könyvkiadó, Budapest. 156–168.
- Gósy Mária – Krepsz Valéria 2017. Szünet a szóban: típusok, jellemzők, időtartamok. In Gósy Mária – Krepsz Valéria: *Morfémák időzítési mintázatai a beszédben*. MTA Nyelvtudományi Intézet, Budapest. 199–225.

- Gyarmathy Dorottya 2017. A néma szünetek funkciói a spontán beszédben. *Beszédkutatás* 2017. 67–92.
- Hegedűs, Lajos 1953. On the problem of the pauses of speech. *Acta Linguistica Hungarica* 3. 1–36.
- Horváth Viktória 2014. *Hezitációs jelenségek a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- Horváth Viktória 2017. Közlések grammatikai szerkesztettsége 6–9 éves gyermekek narratíváiban. *Anyanyelv-pedagógia* 2017/4. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=703>
- Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- Ingram, Jenni – Elliott, Victoria 2014. Turn taking and ‘wait time’ in classroom interactions. *Journal of Pragmatics* 62. 1–12.
- Kowal, Sabine – O’Connell, Daniel C. – Sabin, Edward J. 1975. Development of temporal patterning and vocal hesitations. *Journal of Psycholinguistic Research* 4/3. 195–207.
- Kurzon, Dennis 2007. Towards a typology of silence. *Journal of Pragmatics* 39. 1673–1688.
- Laczkó Mária 2009. Középiskolai tanulók spontán beszédének temporális jellegzetességei. *Magyar Nyelvőr* 133/4. 447–467.
- Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).
- Lounsbury, Floyd G. 1965. Transitional probability, linguistic structure and system of habit-family hierarchies. In Osgood, Charles E. – Sebeok, Thomas A. (eds.): *Psycholinguistics. A survey of theory and research problems*. Indiana University Press, Bloomington–London. 93–101.
- Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás* 2005. 63–77.
- Maroni, Barbara 2011. Pauses, gaps and wait time in classroom interaction in primary schools. *Journal of Pragmatics* 43. 2081–2093.
- Menyhárt Krisztina 2003. A spontán beszéd megakadásjelenségei az életkor függvényében. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia a gyakorlatban*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen. 125–138.
- Mészáros Katalin Edit 2012. *Megakadásjelenségek és a néma szünetek – kategóriák összefüggései*. Első Század online. 11. évfolyam 2. szám. [http://epa.oszk.hu/01600/01639/00006/pdf/EPA01639\\_első\\_szazad\\_2012\\_nyar\\_12\\_1-145.pdf](http://epa.oszk.hu/01600/01639/00006/pdf/EPA01639_első_szazad_2012_nyar_12_1-145.pdf)
- Neuberger Tilda 2014. *A spontán beszéd sajátosságai gyermekkorban*. ELTE Eötvös Kiadó, Budapest.
- Olaszy Gábor 2005. Prozáiai szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás* 2005. 21–50.
- Ramanarayanan, Vikram – Bresch, Erik – Byrd, Dani – Goldstein, Louis – Narayanan, Shrikanth S. 2009. “Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation,” *Journal of Acoustical Society of America* 126. 160–165.

- Sacks, Harvey – Schegloff, Emanuel A. – Jefferson, Gail 1974. A simplest systematics for the organization of turn taking for conversation. *Language* 50. 696–735.
- Sallai János – Szende Tamás 1995. Spontán közlések beszédpszüneteinek pszicholingvisztikai értelmezése (egészséges és skizofrén közlők szövegeiben). *Általános Nyelvészeti Tanulmányok XVIII.* 209–222.
- Schönpflug, Ute 2008. Pauses in elementary school children's verbatim and gist free recall of a story. *Cognitive Development* 23/3. 385–394.
- Singh, Latika – Shantisudha, P. – Singh, Nandini C. 2007. Developmental patterns of speech production in children. *Applied Acoustics* 68. 260–269.
- Sweet, Henry 1890. *A primer of phonetics*. Clarendon Press, Oxford.
- Szabó Kalliopé 2008. Megakadási jelenségek nyolcévesek spontán beszédében. *Anyanyelv-pedagógia*. 2008/2. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=56>
- Szende Tamás 1979. A szünet és a junktúra. *Magyar Fonetikai Füzetek* 4. 7–32.
- Vakula Tímea 2012. Diszlexiás gyermekek spontán beszédének vizsgálata. In Navracsecs Judit – Szabó Dániel (szerk.): *Mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III.* Tinta Könyvkiadó, Budapest. 199–207.
- Vakula Tímea – Krepsz Valéria 2017. *Egyéni sajátosságok vizsgálata a spontán beszéd temporális szerkezetében*. Előadás. Elhangzott: A nyelv közösségi perspektívája. Nagyvárad. 2017. június 30.
- Váradi Tamás 1988. A beszédpszünet szubjektív és objektív regisztrálásának összevetéséről In Kontra Miklós (szerk.): *Beszélt-nyelvi tanulmányok*. MTA Nyelvtudományi Intézet, Budapest. 44–59.
- Váradi Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás 2010*. 100–109.
- Zellner, Brigitte 1994. Pauses and the temporal structure of speech. In Keller, E. (ed.): *Fundamentals of speech synthesis and speech recognition*. John Wiley, Chichester. 41–62.

### Köszönetnyilvánítás

A kutatást a Bolyai János Kutatási Ösztöndíj támogatta.

### Functions of silent pauses in children's spontaneous speech

Silent pauses have several functions in spontaneous speech (e.g. breathing, marking syntactic boundaries as well as speech planning difficulties, time for self-repair). The aim of the study is the analysis of the interrelation between the temporal pattern and the syntactical position of silent pauses in Hungarian children's spontaneous speech. Our hypotheses are the following: (i) silent pauses show different temporal characteristics depending on their position, (ii) different temporal patterns would be proved in children's pauses depending on their age. 40 narratives were recorded with 6, 7, 8- and 9-year-old children. The total material was about two hours long (4 minutes per participants, on average) and was annotated using Praat 5.1. The frequency,

the position and the duration of silent pauses were analyzed. Statistical analysis was carried out using SPSS 20.0 program. We differentiated silent pauses (S) occurring either in phrases or at phrase boundaries and pauses functioning as editing phases (E). Results showed that the position and function of silent pauses had a main effect on both their occurrences and their durations in the children's speech similar to the adults' speech. However, no differences could be pointed out on the durations of silent pauses between the age groups.

## A DICHOTIKUS SZÓÉSZLELÉS ÉLETKOR- SPECIFIKUS JELLEMZŐI

**Krepsz Valéria – Gósy Mária**

MTA Nyelvtudományi Intézet

### **Bevezetés**

Az emberi agy egyik jellemzője, hogy a két agyfélteke mind morfológiai-  
lag, mind funkcionálisan aszimmetrikus (Kandel et al. 2000; Toga–Thompson 2003; Halpern et al. 2005; Riès et al. 2016). A jobb és a bal agyfélteke működése között sajátos funkciómegosztás figyelhető meg, ez azt jelenti, hogy a két félteke funkcióik szerint elkülönülve, de folyamatos együttműködésben végzik a vezérlést (pl. Brobäck 1988; Garman 1990). Mindennek evolúciós előnyei, például a zsákmányszerzés során megosztott figyelem, már régóta ismertek (pl. Bisazza et al. 1998; Halpern et al. 2005; Doron–Gazzaniga 2008).

A bal agyfélteke felel például a nyelvi folyamatok működtetéséért, az írásért, elemzései reálisak, logikus gondolkodás jellemzi, feladata az algebrai feladatok elvégzése. Noha a nyelvi feladatok jóval nagyobb mértékben igénylik a bal agyféltekei működést, a jobb agyfélteke is részt vesz ezekben (Tervaniemi–Hugdahl 2003). A jobb agyfélteke elsősorban képekben gondolkodik, itt található az érzelmek központja; kreativitás, játékoság és fantázia jellemzi (pl. Baddeley 2001; Steinberg et al. 2001; Gósy 2005; Kállai 2008). Ez azonban nem jelenti azt, hogy egy-egy képesség vagy tulajdonság kizárólagosan az egyik oldalhoz tartozna. A hippocampus például alapvetően a memória-folyamatokban játszik szerepet: a baloldal inkább a verbális, a jobb oldal inkább a téri-vizuális memóriaműködésben tölt be nagyobb szerepet. Noha a bal agyféltekében található a beszéd központja (legalábbis az emberek igen nagy többségénél), az új, avagy ritkán használt szavak esetében a jobb, míg a jól ismert szavaknál a bal agyfélteke mutat nagyobb aktivitást (Tervaniemi–Hugdahl 2003).

Az agyféltekék működésére tehát jellemző a funkciómegosztás és az együttműködés, mindez úgy megy végbe, hogy az egyik agyfélteke vezető szerepet tölt be a vezérlésben, vagyis domináns. Ezt nevezik agyfélteke-dominanciának vagy agyi lateralizációnak. Az agyfélteke-dominancia az anyanyelv-elsajátítás során, a különböző nyelvi ingerek hatására alakul ki, azonban a habituációs vizsgálatok azt mutatták, hogy már a csecsemőknél is igazolható bizonyos dominanciák jelenléte (Pléh 1981; Hugdahl 2003; Gósy

2005). Az agyféltekei dominancia többféle módszerrel is kimutatható (Bethmann et al. 2007).

Geschwind és Galaburda (1987) komplex modelljében az agyi lateralizáció kialakulását elsősorban a tesztoszteronszint magzatkori emelkedésével magyarázzák, amely véleményük szerint hatással van az idegi fejlődésre, az immunrendszer kialakulására, valamint a velőcső fejlődésére is. Ebből adódóan összefüggést feltételeznek a cerebrális lateralizáció, valamint a tanulási nehézségek, a „tehetség” és az immunrendszerrel kapcsolatos problémák között. Szintén a hormonháztartás és az agyi dominancia kialakulásának kapcsolatával magyarázzák a nemek között megfigyelhető, a beszédben, illetve a nyelvhasználatban megjelenő különbséget is. A feltételezések szerint ugyanis a fiúknál mérhető magasabb tesztoszteron-koncentráció lassítja a bal agyfélteke, ezáltal pedig a beszédcentrum idegsejtjeinek differenciálódását, amelyből adódóan a fiúknál bizonyos jobb agyféltekés sajátosságok erőteljesebb fejlettsége és jobb működése figyelhető meg (Steinberg et al. 2001). A neurofiziológiai tanúság, csakúgy, mint a strukturális, illetve a funkcionális képalkotó eljárások eredményei azt mutatják, hogy a jobbkezes emberek túlnyomó többségénél a bal agyfélteke domináns a beszéd folyamataiban (Penna et al. 2007; Hakvoort et al. 2016). Az agyfélteke-dominancia összefüggést mutat a kezességgel, és jellemző az egyénre (Bryden 1970; Hámori 1999). A beszédpercepció kurrens modelljei a bal agyfélteke meghatározott hullámainak azonosításán alapszanak (Hagoort–Indefrey 2014).

A lateralizáció vizsgálatának tradicionális, nem-invazív módszere a **dichotikus tesztek** végzése (Bryden 1988; Meyers et al. 2002; Hugdahl 2003; Thomsen et al. 2004; Fernandes et al. 2006; Helland et al. 2008; Obrzut–Mahoney 2011). Broadbent (1954) volt az első, aki kísérletében különböző szavakat juttatott a résztvevők fülébe, és megállapította, hogy a többségük a jobb fülbe érkezőkből tudott többet helyesen visszamondatni. Doreen Kimura (1961) írta le elsőként ennek a fiziológiai hátterét. A dichotikus eljárás történetét részletesen ismerteti a Jerger–Martin szerzőpáros (2004) és Kenneth Hugdahl 2011-es tanulmányában. Számos tanulmány, amelyek különféle módszereket alkalmaztak (pl. PET, fMRI, MEG, elektrofiziológiai mérések) igazolta a jobbfül-fölény anatómiai alapját (pl. Hugdahl et al. 1999; Brancucci et al. 2005; Penna et al. 2007).

A dichotikus tesztek lényege az, hogy a két fülbe egyidejűleg különböző nyelvi információt (pl. szavakat, hangkapcsolatokat) közvetítenek, és a kísérleti személynek a megadott instrukciókat követve kell hangosan elismételnie a hallottakat (pl. Moncrieff–Wilson 2009). A visszamondott szavak alapján állapítható meg a fülfölény (‘ear advantage’), amely utal az agyfélteke-dominanciára. A jobbfül-fölény (angol rövidítése: REA, azaz ‘right-ear advantage’) azt jelenti, hogy a kísérleti személy több elhangzott nyelvi egységet tudott megismételni a jobb fülébe érkezettek közül, ami arra utal, hogy a bal agyféltekéje a domináns. A balfül-fölény (angol rövidítése: LEA, azaz

'left-ear advantage') a jobb agyfélteke domináns voltára utal. Ha nem mutatható ki fülfölény (angol rövidítése: NEA, azaz 'no-ear advantage'), az bilaterális vagy keverten működő dominanciát feltételez, de jelenthet bal temporális diszfunkciót is. Az, hogy valakinél a dichotikus teszt alapján bal vagy jobb agyfélteke-dominanciát állapítottak meg, megegyezett a Wada-teszt eredményével (Hugdahl et al. 1997).

A módszer alkalmazása az elmúlt közel 60–70 évben elterjedté vált (Hugdahl 2011; Obrzut–Mahoney 2011), a központi auditív feldolgozás egyik vizsgálati eljárása. A nyelvi ingerek a feladat típusától függően változnak, a cél lehet a binaurális információk integrációjának, avagy a binaurális információ szegregációjának a vizsgálata (Moncrieff et al. 2013). A tesztek anyaga a szótagoktól a szavakon, a számokon, illetve az értelmetlen hangsorokon át a mondatokig változik (pl. Kimura 1961; Willeford 1977; Musiek 1983; Meyers et al. 2002; Bethmann et al. 2007; Moncrieff 2011; Andrade et al. 2015). A gyakoriságot tekintve a dichotikus tesztek többsége szótagokat tartalmaz (leggyakrabban CV típusúakat, pl. Hugdahl 1995; Sætrevik 2012). Olykor a fülbe érkező nyelvi ingerek intenzitása is változó (Penna et al. 2006).

A dichotikus tesztek eredményei igazolták a jobbful-fölényt az átlagos, egészséges populációban (Lebel–Beaulieu 2009; Ettinger–Veenstra et al. 2012; Dawes–Bishop 2010; stb.). A nyelvi és módszertani különbségek azonban befolyásolják a fülfölény, valamint az agyfélteke-dominancia megállapítását, illetőleg annak mértékét. Azok a tesztek, amelyekben számokat használtak, kisebb mértékű dominanciát mutattak ki (pl. Kimura 1961; Strouse–Wilson 1999), mint azok, amelyek CV típusú szótagokat (pl. Hugdahl 1995) vagy egy szótagból álló szavakat tartalmaztak (Strouse–Wilson 2001). Moncrieff egyik kutatásában azt találta, hogy az életkornak nem volt szignifikáns hatása a teljesítményre a szavak esetében, de a számok alkalmazásakor igen (2011).

Az eredményeket nyilvánvalóan befolyásolja az instrukció is, amely lehet ún. 'szabad visszamondás' ('free recall'), de a kísérletvezető irányíthatja is a résztvevő figyelmét az egyik vagy a másik fülbe érkező nyelvi ingerekre, avagy az elvárt szavak mennyiségére. Moncrieff igazolta azt is, hogy az adatok számolásának (számításának) módja meghatározó tényezője az eredménynek (2011). Noha ezek a módszertani különbségek megnehezítik, olykor ellehetetlenítik a különböző kísérletek adatainak összevetését, a dichotikus teszt kiválóan alkalmas az agyféltekei aszimmetria, illetve a beszéd-lateralitás kimutatására, sőt magasabb kognitív funkciók megismerésére is (Studdert-Kennedy–Shankweiler 1970; Hugdahl 2011). Magyarai közel ötven évvel ezelőtt a centrális hallópályák működését vizsgálta dichotikus teszttel magyar anyanyelvű felnőttek részvételével (1970).

Az agyfélteke-dominancia kialakulásának, illetve kimutathatóságának időpontjáról sokszor jelentősen ellenmondó adatokat közölnek a kutatások. A szerzők egy része feltételezi, hogy a tipikus fejlődésű gyermekek esetében létezik egy kritikus időszak, ameddig a dominanciának ki kell alakulnia, ez a 6-7 éves kor (pl. Kimura 1961). Az adatok szerint a bal féltekei dominancia, illetve a jobbfül-fölény nyelvi ingereknél 5-7 év körül alakul ki, és 10 éves kor táján szilárdul meg. Agyi aszimmetriákat azonban az észlelési vizsgálatok már újszülött kortól igazoltak; 3 éves gyermekeknél ki lehetett mutatni dichotikus teszt alkalmazásával (Pléh 1981; lásd még: Pléh–Lukács /szerk./ 2014). Krashen (1973) kutatásaiban a dominancia kialakulását a formális gondolkodás megszilárdulásával hozta kapcsolatba, így a lateralizáció kritikus periódusának határát a pubertáskorhoz köti. Hatéves kor után, illetve iskolásoknál egyes kutatások nem mutattak ki változást az életkor előrehaladtával, míg mások igen (vö. Obrzut–Mahoney 2011; Moncrieff 2011; illetve Bryden 1970; Carlsson et al. 2011).

A két agyfélteke közötti kapcsolat fejlődésében 6 éves kortól igazolták egy olyan meghatározó időszak kezdetét, amikor a kérgestest fiziológiai és funkcionális fejlődése megindul (Westerhausen et al. 2011). A szerzők hat- és nyolcéves gyermekekkel végzett vizsgálataikban azt találták, hogy a két agyfélteke közötti információáramlás (a dichotikus tesztek eredményei alapján) és az adatolt fiziológiai különbségek szoros összefüggést mutatnak; a fejlődés éppen ebben az időszakban mutatható ki. Megerősítették, hogy a bal fülbe érkezett szótagok helyes felismerése szoros kapcsolatban van a kérgestest fejlettségével. Úgy gondolják továbbá, hogy nem véletlen, hogy ebben az életkorban stabilizálódik a fonológiai tudatosság is, amely kapcsolatban lehet mind a fiziológiai, mind a funkcionális változásokkal.

Moncrieff (2011) öt és tizenkét év közötti gyermekek teljesítményét vizsgálta dichotikus tesztekkel, amelyekben egy szótagú szavak, valamint számok alkották a hangingereket. Az eredmények szerint jobbfül-fölényt az 5–7 éves gyermekek közel 60%-a, a 8–10 évesek több mint 75%-a, míg a 11–12 évesek mintegy 70%-a mutatott. A balfül-fölény a következőképpen alakult a három korcsoportban. A legfiatalabbaknál közel 30%-ban, a 8–10 éveseknél több, mint 20%-ban, a legidősebbeknél pedig kicsit több mint 25%-ban volt kimutatható. Egy másik kutatásban 9 évesek és 11-12 évesek dichotikus tesztben nyújtott teljesítményét elemezték (Hakvoort et al. 2016). A teszt a fiatalabbaknál számokat, az idősebbeknél CV szótagokat tartalmazott. Az eredmények mindkét korcsoportban jobbfül-fölényt igazoltak (a 9 éveseknél kb. 63%-ban szemben az idősebbek mintegy 60%-os előfordulásával). A két fülbe közvetített nyelvi ingerek közötti különbség kissé növekedett az életkor előrehaladtával.

Moncrieff és Musiek kutatásában (2002) 11 évesek helyes visszamondásait elemezték dichotikus tesztekben, amelyek nyelvi anyaga számokból, szavakból és CV típusú szótagokból állt. A jobb fülbe érkezett szótagok helyes fel-



ismerésének átlaga 55,9%-nak, a bal fülbe érkezőké pedig 46,9%-nak adódott. A számok helyes ismétléseinek arányai a jobb fül esetén 92,2%-nak, a bal fül esetén 94,3%-nak adódtak. A szavakat tartalmazó dichotikus tesztben a jobb fülbe érkezettek helyes visszamondása 88%, a bal fülbe érkezetteké pedig 82% volt. (vö. 432. oldal.) A számok alkalmazása esetén nem volt szignifikáns különbség a két fül tekintetében, a szavak és a CV szótagok esetében azonban igen.

Számokat tartalmazó dichotikus tesztet vettek fel 200 gyermekkel 5 és 13 éves kor között (Rosenberg 2011). Két-két számot hallottak a résztvevők a jobb oldali és a bal fülükbe, amelyeket ismételnük kellett. A helyes visszamondások növekedést mutattak az életkor mentén, a két fiatalabb korcsoportban a jobb fülbe érkezők voltak nagyobb mértékben helyesek; a 11-12 éveseknél gyakorlatilag nem volt különbség. Az 5-6 évesek átlaga a bal fülbe érkező számok esetén 52,5%, a jobb fülbe érkezők esetén 69,9% volt. A 7-8 évesekéi a bal fülben hallottak esetén 61,3%, a jobb fülben hallottaknál 73,9%, míg a 11-12 évesekéi a jobb fülbe közvetítetteknél 92,6%, a bal fülben hallottak esetén 88,1% volt.

Keresztmetszeti vizsgálatban elemezték (11 évestől felnőttekig), hogy vajon az egyidejűleg közvetített egy, két, illetve három számot miként tudják a résztvevők visszamondani a dichotikus teszt helyzetben (Moncrieff–Wilson 2009). A bal fülbe hallott szavak minden korcsoportban átlagosan kisebb mértékben voltak ismételhetők, mint a jobb fülben hallottak, továbbá ahogy növekedett a hallott szavak száma, úgy csökkent a helyes felismerés. A vizsgált életkori csoportok átlagai közel azonosak voltak. Az átlagokat tekintve nagymértékű jobb-fül-fölényt mutattak ki minden korcsoportban. A két legfiatalabb csoport átlagos adatai az alábbiak voltak. A számpárok helyes ismétlése a 10–11 éveseknél a bal fül esetében 95,4%, a jobb fül esetén 98,6% volt; két-két szám esetén 84,2%, illetve 92,3%. A 12–14 éveseknél a számpárok esetén a bal fülbe érkezők helyes ismétlésének aránya 96,9%, a jobb fülbe érkezőké pedig 97,9% volt. A két-két számpár esetén az arányok a következőképpen alakultak: 85,8%, illetve 93,6%. A gyermekeknél – a fiatal felnőttekhez képest – a fül-fölény nagyobb mértékű, amit a szerzők a figyelem és a munkamemória fejlődésével magyaráznak (Moncrieff–Wilson 2009).

Iskolásoknál a dichotikus tesztekkel kapott eredmények általában megerősítették a bal agyfélteke dominanciáját, valamint azt, hogy hatéves kor után nem volt változás kimutatható a lateralizációban (Bryden–Allard 1981; Asbjørnsen–Helland 2006; Obrzut–Mahoney 2011; Moncrieff 2011). Más kutatások ellentmondó eredményre jutottak, 9-, 13- és 17 évesekkel végzett tesztelések szerint az idősebbek pontosabban ismerték fel a nyelvi ingereket a fiatalabbaknál a dichotikus tesztekben (Piazza et al. 1985). Carlsson és munkatársainak tipikus fejlődésű 9 és 14 évesekkel végzett vizsgálatában a gyermekek 20%-ánál nem volt kimutatható agyfélteke-dominancia (2011).

Bryden leírta, hogy a jobbfül-főlény előfordulása növekedett a jobb kezes másodikosoktól a jobbkezes hatodikosokig (1970).

A fülfőlény, illetve az agyfélteke-dominancia kimutatása többféle számolási módszerrel is lehetséges, amelyek alapján meghatározzák az ún. lateralizációs indexet (LI) (pl. Hugdahl et al. 2003; Hakvoort et al. 2016). Ez a mínusz vagy plusz értékű szám utal arra, hogy az adott személynél a dichotikus tesztben a REA, LEA vagy (0 esetén) NEA volt kimutatható. A lateralizációs indexek Hakvoort és munkatársainak kutatásában nem mutattak eltérést a 9 évesek és a 11-12 évesek között; megfogalmazásukban az LI a vizsgált életkorban stabil.

Magyar anyanyelvű óvodásokkal és iskolásokkal, tipikus fejlődésűekkel és beszédfeldolgozási nehézséget, nagyothallást és olvasástanulási zavart mutatókkal végeztek vizsgálatokat (Pléh 1981; Gósy 1996; Reinhardt 2003; Gósy et al. 2018). Az első kísérletsorozatot Pléh Csaba végezte, 3–6 éves óvodásokat vizsgált (1981). Adatai szerint a 4-5 évesek 46–52%-ban mutattak kialakult agyfélteke-dominanciát; a gyermekek mintegy a felénél a jobb agyfélteke volt domináns, és kevesebb, mint 20%-uknál volt kimutatható bal agyféltekei dominancia. Reinhardt (2003) 126 gyermeket vizsgált, 4-5, 7-8 és 9-10 évesek csoportjaiban tesztelte az agyfélteke-dominancia meglétét (ugyanazon teszttel, mint Pléh). Eredményei azt mutatták, hogy a 4-5 éves óvodások 65%-ánál, a 7-8 évesek 40%-ánál, a 9-10 évesek 58%-ánál volt kimutatható dominancia (az óvodásoknál gyakrabban fordult elő, hogy a jobb agyfélteke volt domináns, mint az iskolásoknál, akiknél jellemzően bal agyféltekei dominanciát tapasztaltak). A 8-10 éveseknél a jobb agyféltekei dominancia előfordulása már csak 13%-os volt.

Reinhardt adatai szerint (2003) a 4-5 éves óvodások 8–11 szót ismételték meg helyesen a dichotikus tesztben, a 7-8 éves iskolások 10–15 szót, a 9-10 évesek pedig 15-16-ot (a gyermekek 6%-a 19 szót is képes volt helyesen visszaadni). Gósy és munkatársai kutatásában (2018) a kontroll csoportot képező, átlagosan 8, 9 és 10 éves gyermekek 14-16 szót mondtak vissza az összes lehetséges 20-ból, általában valamivel kevesebbet a bal, és valamivel többet a jobb fülükbe érkezőkből (a teszt szóanyaga azonban különbözött a korábbi vizsgálatokban használttól). Az eredmények a szavak különbözősége ellenére hasonlóak. Az olvasási problémákkal küzdő gyermekek esetében jelentősen alacsonyabb volt a kimutatható agyfélteke-dominancia aránya, és a visszamondott szavak száma is kisebb volt, mint a kontroll csoport esetében.

A szakirodalom viszonylag egységesnek tekinthető a tekintetben, hogy milyen következményekkel járhat az agyfélteke-dominancia nem megfelelő kialakultsága. A tapasztalatok megerősítették, hogy ha a lateralizáció nem alakul ki megfelelően egy adott életkorig, az elmaradásokhoz, zavarokhoz vezet(het) a beszéd (főként a beszédfeldolgozás) fejlődésében (pl. Gósy 1996; Ettinger-Veenstra 2010; Gósy-Gyarmathy 2011; Gósy et al. 2018). Nemzetközi és hazai tanulmányok sora igazolta, hogy a különféle zavarok, például a

specifikus nyelvi zavar, az írott nyelv elsajátításának nehézsége, a diszlexia, a tanulási zavarok, avagy az autizmus mind-mind összefüggést mutatnak a dichotikus tesztben nyújtott gyengébb teljesítménnyel (Gyarmathy 1998; Moncrieff–Musiek 2002; Dlouha et al. 2007; Moncrieff 2010; Billett–Bellis 2011; Obrzut–Mahoney 2011). Gósy 5 és 12 év közötti, atipikus nyelvfejlődést mutató gyermekekkel vett fel szavakat tartalmazó dichotikus tesztet (1996). Adatai az óvodások mintegy 50%-ánál, az iskolások csupán 34%-ánál jeleztek kimutatható agyfélteke-dominanciát.

Mi az, ami az eddigi kutatásokból jellemzőnek látszik a dichotikus tesztekben kapott eredményekre? Megállapítható, hogy a jobbfül-, illetve balfülfölény aránya iskoláskorban nem változik jelentősen, bár a konkrét életkort tekintve nem azonosak a kutatási eredmények. A szakirodalom ellentmondásos eredményeket közöl az agyfélteke-dominancia fejlődésének életkori meghatározottságáról. A két fül közötti különbség csökkenése a helyesen visszamondott szavak számának tekintetében a fejlődést mutatja (Moncrieff 2011). Az alkalmazott módszertan nagymértékben befolyásolja a kapott adatokat. A fülfölény az agyféltekei dominanciára utal, függetlenül attól, hogy milyen számítási módszert alkalmaztak a kimutatásra (pl. Ettinger–Veenstra et al. 2010). Számos kutatás célja elsősorban valamilyen atipikus anyanyelvi fejlődés és különböző típusú tanulási nehézségek kapcsolatának elemzése volt a fülfölény, illetve a lateralizáció tekintetében. Ezek igazolták, hogy a nem tipikus fejlődésű kísérleti személyek nagyobb mértékben mutatnak balfülfölényt, mint a tipikus fejlődésűek, és a teljesítményük a dichotikus tesztben nem éri el a kontroll személyekét. A dominancia kialakulásának késése, illetve nem egyértelmű kimutathatósága, továbbá a dichotikus teszt során helyesen megismételt szavak alacsony száma, a két fül teljesítményének nagy különbsége (függetlenül attól, hogy a jobb vagy a bal fül teljesítménye volt a jobb) mind-mind problémát jelent a működésekben (Moncrieff 2011). Az életkor-specifikus, normatív adatok ugyanakkor a metodológiai eltérések miatt kérdésesek. Nyilvánvaló, hogy a nyelvek közötti különbségek is hozzájárulnak a sokféleséghez. Mindezek ellenére a kutatók egyetértének abban, hogy a dichotikus tesztek eredményei jól jellemzik mind a gyermekeket, mind a felnőtteket, és számos tekintetben felhasználhatók klinikai populációban is.

Minthogy a dichotikus tesztekkel végzett kutatások nagy többsége angol nyelvűekkel történt, ezért Bless és munkatársai 64 különböző anyanyelvet képviselő, 4408 résztvevő adatait dolgozták fel (2015). A dichotikus teszt mobilalkalmazásként futott (*iDichotic*), a teszt anyaga 36 CV szótagpárból állt, és (brit) angol, norvég, német és észt anyanyelvi bemondók ejtésén alapult (Bless et al. 2015). Az átlagos életkor 33 évnek adódott, a legfiatalabb résztvevők 8 évesek voltak (nők és férfiak, illetve jobb- és balkezesek). A számításokban a lateralizációs indexet alkalmazták. Az eredmények mindentől a jobbfül-fülényt igazolták nyelvtől függetlenül. Az anyanyelvi és a nem

anyanyelvi résztvevők eredményei különböztek, azaz a nyelvi háttér befolyásolta az LI alakulását. Az életkornak nem volt hatása a fülfölny alakulására. A jobbkezesek és a férfiak nagyobb mértékű jobbful-fölnyt mutattak, mint a balkezesek, illetve a nők.

Magyar anyanyelvű gyermekek dichotikus tesztben nyújtott teljesítményére vonatkozóan – mint láttuk – vannak sporadikus adatok, olyan rendszerszerű kísérlet eredményei azonban nem állnak rendelkezésre, amelyek az anyanyelv-elsajátítás folyamatában szemléltetnék a két fülben hallott eltérő nyelvi ingerekre adott válaszok életkor-specifikus változásait. Kutatásunk fő kérdése az volt, hogy hogyan alakul a gyermekek auditív-fonetikai feldolgozása a szófelismerésben, ha a két fülbe eltérő beszédingereket juttatunk.

A kutatás célja annak dokumentálása, hogy (i) tipikus beszédfejlődésű, 3 és 10 év közötti, magyar anyanyelvű gyermekek hogyan képesek a két fülből érkező különböző beszédingereket észlelni és integrálni, (ii) kimutatható-e fejlődés abban az értelemben, hogy növekszik a helyesen visszamondott szavak aránya az életkor növekedésével és (iii) a fülfölny milyen megoszlást mutat az egyes életkori csoportokban. Hipotéziseink a következők voltak: (i) A gyermekek fokozatos fejlődést mutatnak a helyesen visszamondott szavak számában 3 és 10 éves kor között, (ii) a jobbful-fölny, és ennek megfelelően a bal agyfélteke-dominancia minden életkorban nagyobb mértékű előfordulást fog mutatni, mint a balfül-fölny, illetve a nem kimutatható fülfölny, és (iii) a dichotikus tesztben nem várunk eltérést a nemek szerint.

### **Módszertan**

A metodológia leírását a kísérleti személyek, majd a dichotikus teszt bemutatásával kezdjük. Ezt követi a tesztfelvétel (a kísérleti eljárás) és az adatfeldolgozás, illetve az alkalmazott statisztikai próba ismertetése.

### **Kísérleti személyek**

A vizsgálatban 320 három és tíz év közötti gyermek vett részt; korosztályonként 40 magyar egynyelvű óvodás és kisiskolás (minden életkori csoportban 20 lány és 20 fiú). Az egyes csoportokban a gyermekek életkora között maximum 3 hónap különbség volt. Például az 5 éveseket azok az óvodások alkották, akiknek a pontos életkora 4;11 és 5;2 közé esett. Hasonlóan, a 9 évesek életkora 8;11 és 9;2 közötti volt. A hatévesek még óvodába jártak. Így 4 óvodás és 4 iskolás csoportot alakítottunk ki.

Valamennyi gyermek ép értelmű és ép halló volt, mindegyikük jobb kezes (a kezességet a GMP-diagnosztika GMP13-as tesztjével ellenőrizték, vö. Gósy 1995/2006). A gyermekek anamnézise szerint (a szülő közölte adatok alapján) egyikük sem volt megkésett beszédfejlődésű, nem küzdöttek sem nyelvi, sem beszédzavarral. Hosszabb ideig nem éltek külföldön. Valamennyi család a budapesti köznyelvet beszélte. A gyermekek hasonló szociokulturális környezetből jöttek.

### A dichotikus teszt

A jelen kutatásban alkalmazott módszer a GMP-diagnosztika 19-es próbájának saját fejlesztésű dichotikus tesztje (Gósy 1995/2006). A teszt első része 10 szót (vagyis 5 szópárt), a második része 10 két-két szóból álló szópárt, összesen 20 szót tartalmazott. A tesztszavakat bemelegítésként 5 szó közvetítése előzte meg, hogy a gyermekek hozzászokjanak a fülhallgatón keresztül történő beszédhez és annak hangerejéhez. Ezek a bemelegítő szavak azonosak voltak a két fülben. Összesen 30 két szótagból álló tesztszót hallottak a gyermekek, 15-öt az egyik, 15-öt a másik fülükbe közvetítve. Példák a szópárokra: *alma/répa*, *tányér/játék*, *banán/pohár*, *csiga/béka*. A teszt szavai a gyermekek számára valószínűsíthetően jól ismert jelentéseket képviseltek.

A szóanyagot férfi felolvasásában rögzítették, a beszélő a hangmagasságát nem változtatta, a szavak ejtése kvázi-monoton volt. A rögzített szavak átlagos időtartama 680 ms volt, az egyes szópárokat 8 mp-nyi néma szünet választotta el. A két-két szópár esetén a köztük lévő szünetet 4 mp-ben határozták meg. A teszt létrehozásakor minden egyes szópárt (kezdetüket és lecsengésüket) számítógépes szoftverrel szinkronizáltak a jobb és a bal csatorna között. Az eljárás során a szavak akusztikai struktúrája nem változott. Három ábra szemlélteti a teszt egy szópárjának (*banán/pohár*) akusztikai lenyomatait. Az 1. ábrán látható a két szó rezgésképe és hangszínképe, a harmadik ábra a két rezgésképet külön-külön, a spektrogram pedig a szavak egyidejű hangzását szemlélteti (1. és 2. ábra).

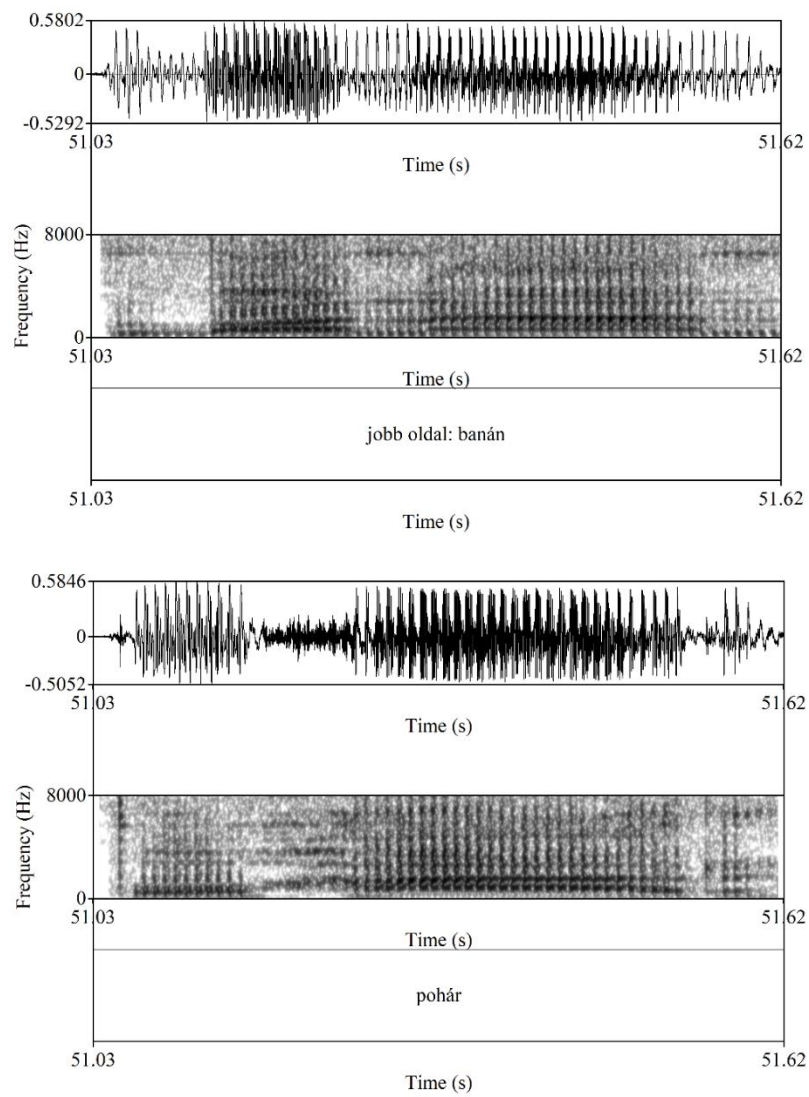
### Tesztfelvétel és adatfeldolgozás

A teszteléskor a gyermek jobb, illetve bal fülébe azonos időben (átlagosan 60 dB-es intenzitáson) közvetítettük a szópárokat fülhallgatón keresztül. Minden felvétel esetében azonos eszközöket használtunk. A gyermekek feladata az elhangzott szavak megismétlése volt szabadon választott sorrendben (ún. szabad visszamondás).

A visszamondott szavakat az elhangzás sorrendjét is jelölve, tesztlapon rögzítettük. Az adatokat négyféle megközelítésben dolgoztuk fel: (i) a helyesen visszamondott szavak számának, illetve arányának elemzése, (ii) az egyszerű kivonás művelete (pl. Moncrieff 2011), (iii) a négyes szópárok figyelembevétele (Bever 1971), (iv) a lateralizációs index számítása (rövidítése: LI), vö. Studdert-Kennedy–Shankweiler 1970).

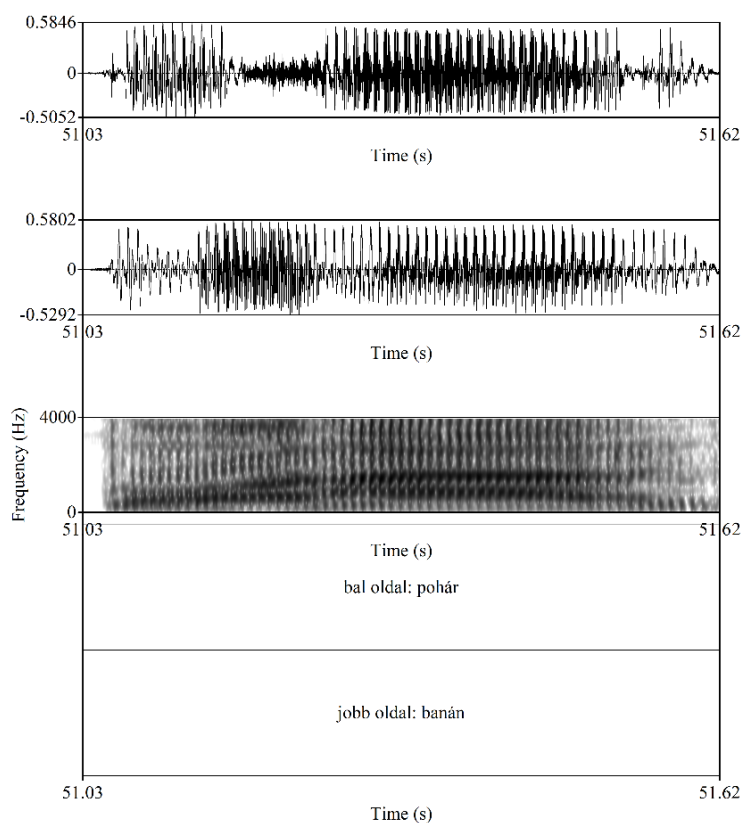
(i) Minden gyermek esetében meghatároztuk a jobb és a bal fülben hallott, helyesen visszamondott szavak számát, illetve összesítettük az összes helyesen ismételt szót. Ennek alapján meghatároztuk a korcsoportokra jellemző teljesítményeket.

(ii) A következő számítás során a jobb fülbe érkező, helyesen ismételt szavak számából kivontuk a bal fülbe közvetített és helyesen visszamondott szavak számát. Ez is utal a fülfölnyire.



1. ábra

A *banán* (fent) és a *pohár* (lent) szavak rezgésképe és hangszíneképe



2. ábra

A *pohár* és a *banán* tesztszavak rezgésképei és az egyszerre hangzás alapján készült hangszínekép

(iii) A négyes szópárok figyelembevételkor az 5x2 szópár, azaz 20 szó helyes visszamondását elemeztük gyermekenként, majd korcsoportonként. A jelölésnek megfelelően a négyes szópárok első párjaiban adatoltuk azokat, amelyeket a gyermek elsőként ismételt. A fülpreferencia kimutathatóságának kritériuma az volt, hogy az egyik fülben hallott szópárok első tagjaiból legalább 4, illetve 5 szó helyes ismétlése megtörténjen.

(iv) Meghatároztuk az ún. lateralizációs indexet (LI, 'laterality index'). Ezt a számítást – úgy véljük – Studdert-Kennedy és Shankweiler publikálta először (1970). A dichotikus teszt anyagát kísérletükben CVC típusú szótagok alkották, ahol a magánhangzókra és a mássalhangzókra eltérő helyes visszamondás volt jellemző attól függően, hogy a résztvevők melyik fülükben hal-

lották azokat. A nullhipotézistől különböző adatok értelmezésére alkalmazták a képletet, és indexnek nevezték (1970: 583), a cikkben azonban azt a szókapcsolatot, hogy 'lateralitási index' még nem használták. A számítási módszert az azóta eltelt évtizedekben széleskörűen alkalmazzák a dominancia kimutatására (Fernandes–Smith 2000; Hugdahl 2003; Hakvoort et al. 2016; stb.). A lateralizációs indexet a dichotikus tesztben helyesen visszamondott szavak (vagy más nyelvi ingerek) alapján a következő képlet szerint számolják ki:  $(\text{jobb fül} - \text{bal fül}) / (\text{jobb fül} + \text{bal fül}) \cdot 100$ . A képletben tehát a 'jobb fül' a jobb fülben hallott és helyesen visszamondott szavak számát, a 'bal fül' pedig a bal fülben hallott és helyesen visszamondott szavak számát reprezentálja. A mínusz értékű LI a jobb, míg a pozitív LI a bal agyfélteke dominanciájára utal.

A statisztikai elemzések során az általános lineáris kevert modellt (GLMM) alkalmaztuk (SPSS 20.0 szoftverben). A függő változók (i) a helyesen visszamondott szavak száma, (ii) a fülfölség és (iii) a lateralizációs index voltak. A független változók pedig (i) a gyermekek életkora és (ii) a fül. A megbízhatósági szintet 95%-ra állítottuk be.

### Eredmények

A dichotikus tesztben kapott adatokat négy alfejezetben mutatjuk be. Az elsőben a tesztszavak helyes észlelésének adatait tárgyaljuk, összesítve, illetve a két fül teljesítménye alapján külön-külön, beleértve a két fül adatainak különbségét is (abszolút értékben). A második alfejezetben a fülpreferencia eredményeit ismertetjük az egyszerű kivonás számításával. A harmadikban a négyes szó párok figyelembevétele történik a fülfölség kimutatásában. A negyedik alfejezetben a lateralizációs index számításának eredményeit ismertetjük.

#### (i) A helyesen visszamondott szavak számának és arányainak elemzése

A vizsgált gyermekek, életkortól függetlenül, átlagosan tíz szót tudtak helyesen megismételni a lejátszott 15-15-ből, azaz 67%-os átlagos teljesítményt értek el. A bal fülben hallottak helyes visszamondásának átlaga a vizsgált populációban 57%, a jobb fül esetében pedig 76%. Az életkor előrehaladtával változnak az eredmények. A két fül adatait összegezve és a fülekbe külön-külön közvetített szavak helyes felismerését tekintve, a helyesen visszamondott szavak száma növekszik a 3 évesek teljesítményétől a 10 évesekéig (1. és 2. táblázat). A kapott adatokat megadjuk a helyesen visszamondott szavak darabszámában és arányában is, így összehasonlíthatók a szakirodalomban különféle képpen található eredményekkel.



1. táblázat: A helyesen visszamondott szavak száma a beszélők életkora és a fülek szerint (összes szó 30 db) (á. = átlag, min. = minimum, max. = maximum)

Kor (év)	A helyesen visszamondott szavak adatai (db)								
	Két fül összesen			Bal fül			Jobb fül		
	á.	min.	max.	á.	min.	max.	á.	min.	max.
3	14	10	18	6	2	8	8	3	12
4	16	10	20	6	4	9	9	4	13
5	17	13	24	7	3	10	11	7	14
6	20	13	25	8	5	12	11	8	15
7	21	15	28	9	6	14	12	9	14
8	23	15	30	10	7	13	12	8	15
9	24	19	28	11	9	15	12	9	15
10	25	21	29	12	8	15	13	11	15

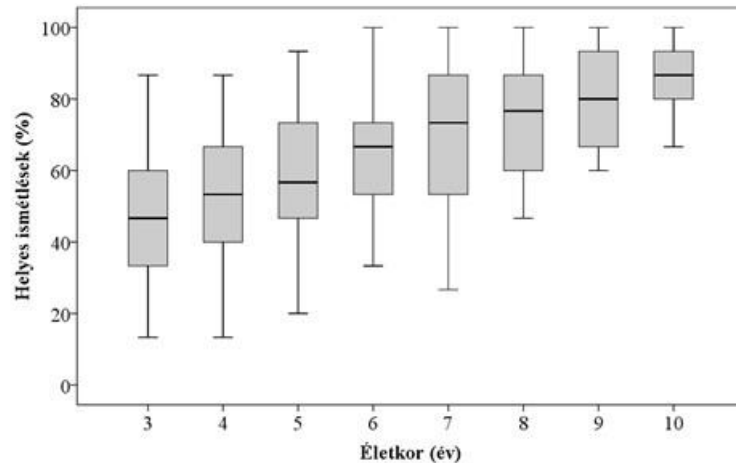
2. táblázat: A helyesen visszamondott szavak arányainak adatai a beszélők életkora szerint

Kor (év)	A helyesen visszamondott szavak adatai (%)					
	Két fül összesen				Jobb fül	Bal fül
	át-lag	átlagos eltérés	mini-mum	maxi-mum		
3	46	21	13	87	53	38
4	52	20	13	87	64	39
5	58	20	13	93	72	43
6	64	17	33	100	77	52
7	68	19	13	93	80	56
8	80	22	7	100	87	72
9	81	14	53	100	88	74
10	85	10	67	100	90	81

Az összes helyes ismétlés számának elemzése (a fültől függetlenül) életkori bontásban igazolja a megfelelően visszamondott szavak számának fokozatos növekedését (3. ábra). Az általunk vizsgált legfiatalabb, hároméves korcsoport adatközlői a lejátszott szavak kevesebb, mint felét ismételték meg pontosan (átlagosan 7 szót). A teljesítmény növekedése 7 éves korig kiegyenlítettnek mondható, átlagosan 6%-os volt, a négyévesek a szavak 52%-át (7,8 szó), az ötévesek a szavak 58%-át (8,7 szó), a hatévesek a szavak 64%-át (9,7 szó) tudták helyesen megismételni. Hat- és hétéves kor között valamivel csökkent ez a növekedés, ekkor csupán 4%-kal növekedett az idősebb gyer-

mekek teljesítménye. Hétéves kortól a növekedés mértéke nagy különbségeket mutatott az egyes korosztályok között. A változás hét- és nyolcéves kor között 12%, a nyolc- és kilencévesek között gyakorlatilag stagnálás tapasztalható, mindkét csoport beszélői átlagosan 12 szót mondtak vissza. A 9 és 10 évesek közötti növekedés mértéke 5%-osnak adódott.

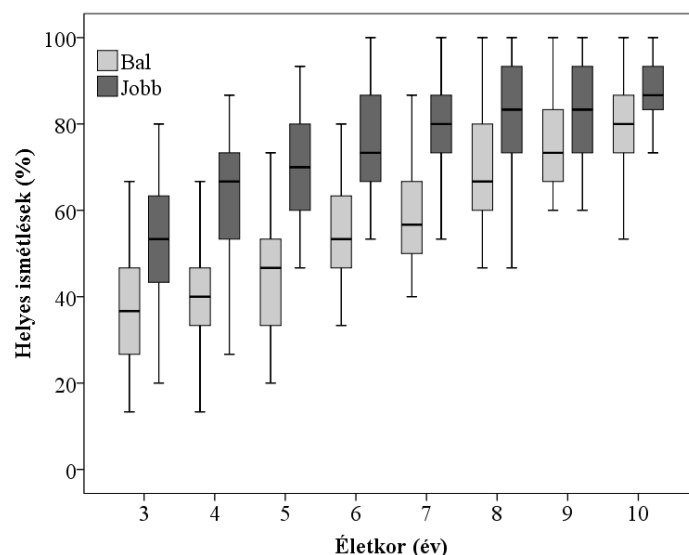
A statisztikai elemzés alátámasztotta a dichotikus tesztben kapott teljesítmények fejlődését 3 és 10 éves kor között. Szignifikáns különbség igazolódott a helyesen visszamondott szavak számában az életkorok mentén [ $F(7, 633) = 37,595$ ;  $p < 0,001$ ]. A nemek között nem volt szignifikáns különbség a helyesen visszamondott szavak számát illetően a csoportokon belül.



3. ábra

A dichotikusan közvetített szópárok helyes ismétlésének aránya a beszélők életkora mentén (medián és szóródás)

Elemeztük a helyesen visszamondott szavak számát, valamint a százalékban kifejezett arányokat aszerint, hogy **a szó a gyermek melyik fülébe érkezett** (4. ábra). A bal fülbe érkezőkből átlagosan két és fél szóval kevesebbet (átlagosan 9 szó, 60%) mondtak vissza a gyermekek, mint a jobb fülbe érkezőkből (itt az átlag 11,5 szó, 76%). A jobb fül preferenciája minden életkori csoport esetében (csoportszinten) kimutatható volt. A legnagyobb különbség a bal és a jobb fül teljesítménye között az ötéves gyermekek esetében jelentkezett, ahol átlagosan 4,5 szóval mondtak vissza többet helyesen a jobb fülbe érkező szavakból, mint a bal fülben hallottakból. A legkisebb különbség a legidősebb, a tízéves gyermekek esetében mutatkozott, itt az eltérés átlagosan 1 szó volt.



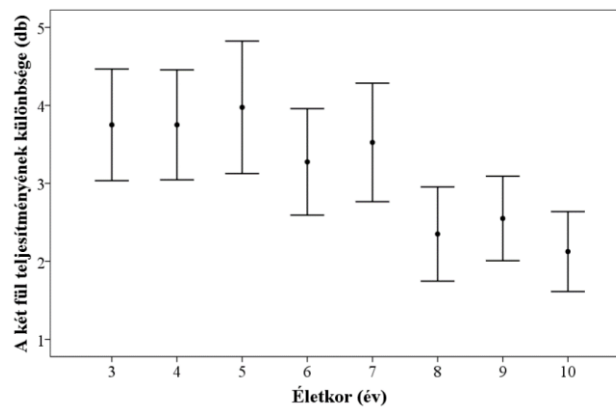
4. ábra

A helyesen visszamondott szavak aránya az életkor és a jobb, illetve bal fül szerinti bontásban (medián és szóródás)

A gyermekek 78,5%-a (251 fő) a jobb, míg 14%-a (45 fő) a bal fülbe érkező szavakból mondott vissza többet; 7,5% (24 fő) esetében pedig nem volt kimutatható fülfőlény a vizsgált csoportokban.

Az életkor előrehaladtával a következő tendencia figyelhető meg: 5 éves korig fokozatosan nő a két fül közötti különbség, majd 5 és 10 éves kor között fokozatosan csökken, azaz közelít egymáshoz a jobb és a bal fül teljesítménye. Az állítás érvényességét azonban óvatosan kell kezelnünk, mivel a jelen vizsgálat egy keresztmetszeti (azaz nem longitudinális) kutatás eredményeit mutatja be. Az 5. ábra a két fülben hallott szavak helyes visszamondásának különbségét szemlélteti a darabszám szerint. Jól látható, hogy a 3–5 évesek jellegzetesen nagyobb különbséget produkáltak a két fül között, a 8–10 évesek pedig jellegzetesen kis különbséget. E tekintetben mintegy a két említett korcsoport között helyezkednek el a 6–7 évesek.

A statisztikai elemzés szignifikáns különbséget igazolt a bal és a jobb fül teljesítménye között [ $F(1, 639) = 107,248$ ;  $p < 0,001$ ], valamint az 'életkor' és a 'fül' interakciója esetében is [ $F(7, 633) = 24,214$ ;  $p = 0,001$ ].



5. ábra

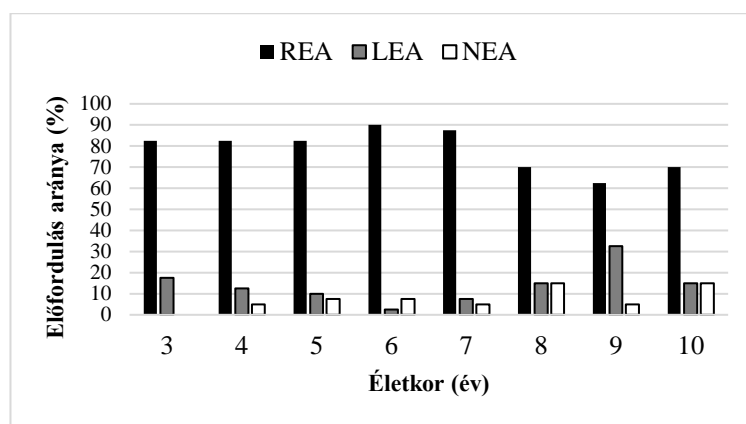
A jobb és a bal fülben hallott szavak helyes visszamondásának különbségei (átlag és szóródás)

#### (ii) Az egyszerű kivonás módszerének alkalmazása

Ebben az elemzésben a jobb oldali fülben hallott, helyesen visszamondott szavak számából kivonjuk a bal oldali fülben hallott, helyesen visszamondott szavak számát (vö. Moncrieff 2011). Ha a jobb fülbe érkezett szavak ismétlése volt nagyobb arányú, akkor pozitív értéket, ha a bal fülbe érkezettek esetén volt nagyobb arányú, akkor negatív értéket kaptunk. A pozitív szám a jobb-fül-fölényre, a negatív szám a balfül-fölényre utalt. A 6. ábra valamennyi gyermek esetén mutatja a fülfölényre kapott adatokat; az ábra elsődlegesen azt a célt szolgálja, hogy a jobb-fül-fölény nagymértékű megjelenését szemléltessük. A kimutatható dominancia a gyermekek 95,7%-ára volt jellemző.

A résztvevők döntő többsége – ennek a számításnak az eredményeként is – jobb-fül-fölényt mutat (a résztvevők 82%-a), a balfül-fölényt mutatók aránya mindössze 13,7%. A fülfölényt nem mutatók aránya pedig 4,3%. Ha korcsoportonként nézzük az adatokat, akkor különbségek figyelhetők meg az életkor függvényében. A 3 évesek 17,5%-a, a 4 évesek 12,5%-a, az 5 évesek 10%-a, a 6 évesek 2,5%-a, a 7 évesek 7,5%-a, a 8 évesek 15%-a, a 9 évesek 32,5%-a, végül a 10 évesek 15%-a mutatott balfül-fölényt. A legnagyobb arányú jobb-fül-fölényt a 6 és a 7 éveseknél adatoltuk. Néhány gyermeknél előfordult, hogy azonos számú szót ismételték helyesen a jobb és a bal fülükben hallottak közül. Az 5 és a 6 évesek csoportjában 3-3 gyermeknél, a 4, a 7 és a 8 évesek esetében 2-2 gyermeknél, a 9 és a 10 évesek között pedig 1-1 gyermeknél tapasztaltuk a NEA előfordulását. Egyáltalán nem fordult elő a 3 évesek csoportjában.

A statisztikai elemzések szignifikáns különbséget igazoltak a jobb és a bal fülbe érkező szavak helyes ismétlése között: [ $F(1, 639) = 9,734; p < 0,001$ ].

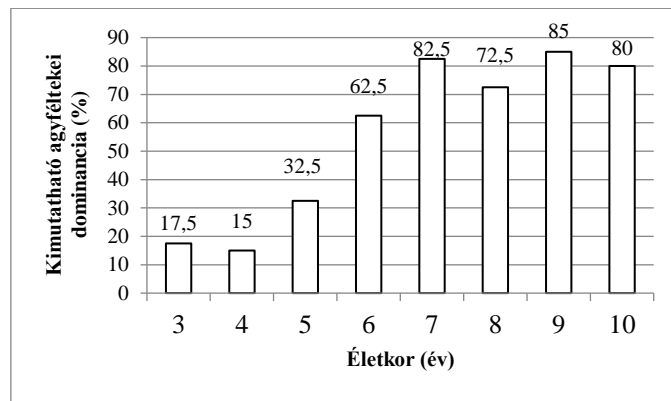


6. ábra

A REA (jobbfül-fölény), a LEA (balfül-fölény) és a NEA (nincs kimutatható fül-fölény) előfordulása a vizsgált korcsoportokban az egyszerű kivonás alkalmazása alapján

### (iii) A négyes szópárok figyelembevétele

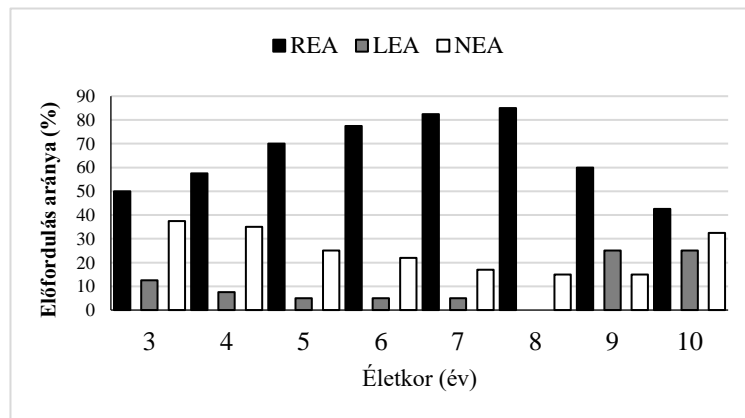
Ennek a számítási módnak az alkalmazásával az eredmények eltértek a fentebbi, egyszerű kalkuláció adataitól. Itt csak a négyes (kettő-kettő) szópárokat vettük figyelembe, közülük is az elsőként elhangzottakat (l. részletebben: Gósy 1995/2006). A vizsgált gyermekek 56%-ánál volt kimutatható valamilyen fül-fölény, illetve agyfélteke-dominancia. Ez a részesedés az életkor előrehaladtával nőtt, a legkisebb arányt (a gyermekek 15%-a) a négyéves, a legnagyobbat (a gyermekek 85%-a) a kilencévesek csoportjában adatoltuk (7. ábra). A legnagyobb növekedés (mintegy 30%) a tekintetben, hogy kimutatható a dominancia, az öt- és hatéves kor között volt megfigyelhető. Ez összefüggést mutat a számos külföldi kutatásban feltételezett kritikus periódus életkori határával. A statisztikai elemzés az életkor mentén szignifikáns különbséget igazolt a kimutatható fül-fölény, illetve az agyféltekék dominanciája között [ $F(2, 318) = 13,253$ ;  $p = 0,004$ ]. A nemek között nem igazolódott szignifikáns különbség.



7. ábra

A kimutatható agyféltekei dominancia aránya a beszélők életkora mentén

A balfül-fölény a gyermekek 17%-ára, a jobbfül-fölény pedig 71%-ukra volt jellemző. Az életkori megoszlást a 8. ábra szemlélteti, amelyen jól látható, hogy minden korcsoportban a gyermekek döntő többsége jobbfül-fölényt, illetőleg bal agyfélteke-dominanciát mutat.



8. ábra

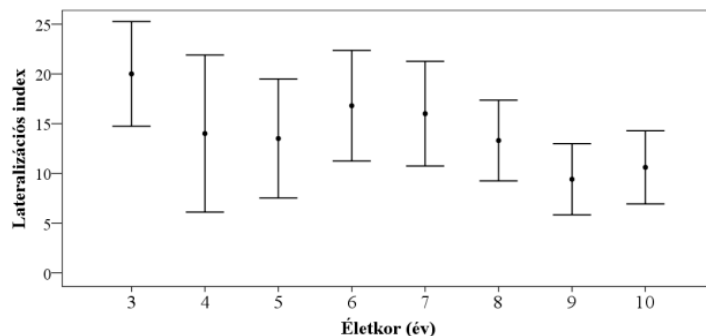
A REA (jobbfül-fölény), a LEA (balfül-fölény) és a NEA (nincs kimutatható fül-fölény) előfordulása a vizsgált korcsoportokban a négyes szópárok figyelembevétele alapján

A legtöbb, jobbfül-fölényt mutató gyermek 8 éves volt, ugyanezt legkevésbé az 10 éveseknél adatoltuk. A bal- és a jobb agyféltekei dominancia

arányának legnagyobb különbsége 7 éves korban volt kimutatható, ekkor a jobb félteke a gyermekek 12%-ánál, a bal félteke pedig 88%-uknál volt dominánsnak tekinthető. Ezt a statisztikai elemzések is alátámasztották. A jobb-fül-főlényt mutatók esetében nem, míg a bal-fül-főlényt mutatók előfordulása szignifikáns különbséget igazolt az életkorok mentén [ $F(2, 318) = 10,143$ ;  $p = 0,012$ ].

#### (iv) A lateralizációs index számítása

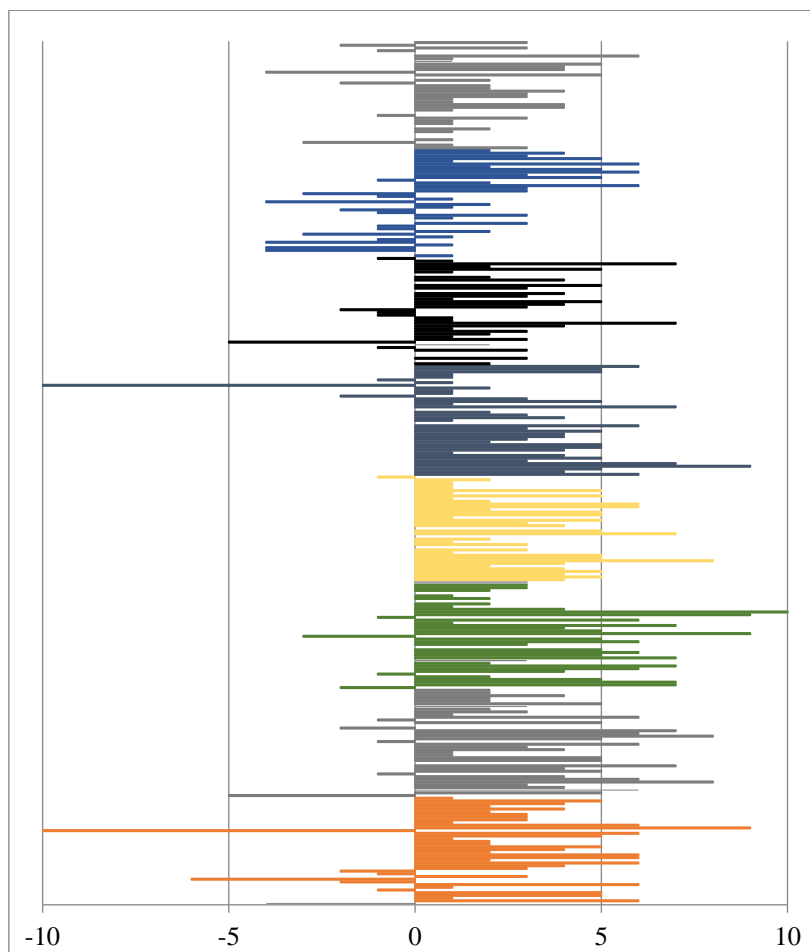
Adatainkon elvégeztük az LI-re vonatkozó számításokat; az eredményeket a 9. ábra, korcsoportonként pedig a 10. ábra szemlélteti (a korcsoportok elkülönítését a színezések segítik). A kimutatható dominancia a gyermekek 83 %-ára volt jellemző. Minthogy a vizsgált gyermekek túlnyomó többsége jobb-fül-főlényt mutatott, így értelemszerűen a csoportszintű LI-k pozitív értékűek lettek.



9. ábra

A lateralizációs index adatai életkoronként (átlag és szóródás)

A lateralizációs index jelentős különbségeket mutat az iskolások és az óvodások között; a nagyobb csökkenés a nyolcéveseknél figyelhető meg. Nyolcéves korig relatíve nagy érték az LI, míg a három legidősebb vizsgált korcsoportban relatíve kicsi LI-értékeket kaptunk. A statisztikai elemzések szerint szignifikáns különbség van a bal és a jobb fülbe érkező szavak helyes visszamondása között [ $F(1, 639) = 11,234$ ;  $p < 0,001$ ].



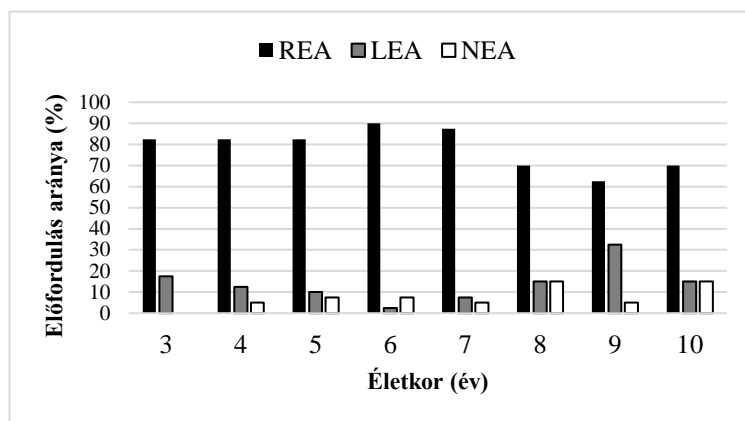
10. ábra

Az LI szerinti fül-főlény alakulása 3–10 éves kor között (a függőleges tengely a gyermekeket, a vízszintes az LI-értékeket szemlélteti (a színek a különböző életkorokra utalnak, a 3 évesek vannak legfelül)

A lateralizációs index alapján is azt tapasztaltuk, hogy a legnagyobb mértékben minden életkorban a REA (jobb-fül-főlény) volt igazolható. Lényegesen kisebb a LEA (balfül-főlény) előfordulása; a 9 éveseknél kiemelkedően nagy volt a megjelenése (11. ábra). Az LI értékeit tekintetbe véve az mondható, hogy a LEA a 6 és a 7 éveseknél fordul elő a legritkábban, mind a fiata-



labbb, mind az idősebb gyermekekénél gyakoribb. Itt is látható, hogy a 3 éveseknél nincs NEA, és az értékek a 9 évesek kivételével enyhén növekvő tendenciát mutatnak az életkor előrehaladtával.

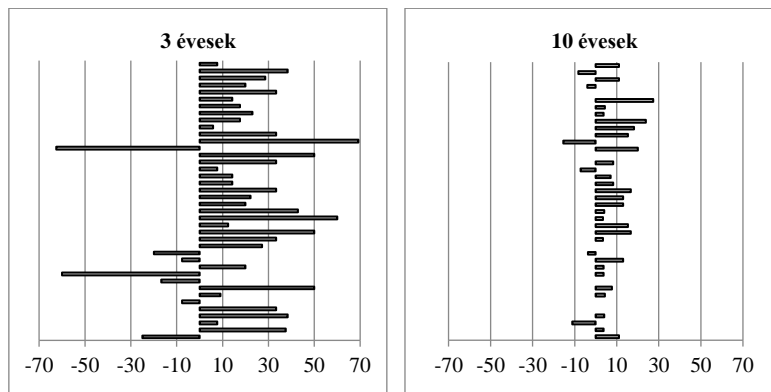


11. ábra

A REA (jobbfül-fölény), a LEA (balfül-fölény) és a NEA (nincs kimutatható fülfölény) előfordulása a vizsgált korcsoportokban az LI alapján

A statisztikai elemzések szerint az LI értékét tekintve az életkor függvényében szignifikáns a különbség [ $F(8, 312) = 5,642$ ;  $p = 0,023$ ]. Példaként szemléltetjük a 3 évesek és a 10 évesek csoportjának összes adatával a pozitív, a negatív és a zéró LI előfordulásait (12. ábra).

Összevetettük a fülfölény kimutatására alkalmazott három különböző számítási módszer adatait. A REA és a LEA előfordulásának arányai a vizsgált korcsoportokban gyakorlatilag nem különböztek az egyszerű kivonás (jobb fül–bal fül) eredménye és a lateralizációs index alapján megállapított dominanciák között. Elterést tapasztaltunk a négyes szópárok figyelembevétele esetén. Itt ugyanis jellegzetesen különböznek az arányok a 3–6 éves és a 10 éves korcsoportokban. A 7, 8 és 9 éves esetében az előfordulások vagy azonosak, vagy kissé eltérően különböznek. Az eltérések minden esetben azt mutatják, hogy a négyes szópárok figyelembevétele alapján végzett számítások esetén az adott életkorokban kisebb arányú a kimutatható dominancia.



12. ábra

A lateralizációs index értékeinek alakulása 3 éves és 10 éves gyermekeknél (a függőleges tengelyeken a negyven gyermeket szemléltettük, a vízszintes tengely az LI értékeit mutatja)

### Következtetések

Korábbi kutatások szerint az agyi funkciómegosztásból adódó agyi aszimmetria, valamint a nyelvi szempontból domináns agyfélteke már hároméves kortól kimutatható (Best 1984; Hiscock 1988; Pléh–Lukács /szerk./ 2014). Az egyik alapkérdés, amelyre ellentmondó válaszokat találunk a szakirodalomban az, hogy vajon a fülfölény, illetve a domináns agyfélteke kimutathatósága mikorra szilárdul meg, és van-e fejlődés az életkor növekedésével. A jelen kutatásban 320 tipikus fejlődésű, három- és tízéves kor közötti gyermeket vizsgáltunk dichotikus teszteljárás segítségével. Noha az adatokat különböző számítások mentén dolgoztuk föl, az eltérő módszerek igen hasonló eredményeket igazoltak.

Az idősebb gyermekek több szót mondtak pontosan vissza, mint a fiatalabban, amely a percepció kontroll, a koordináció és a két fülből jövő információ integrálásának fejlődésére utal. Tekintetbe kell venni továbbá a memória kapacitásának növekedését, ami nyilvánvalóan hatással van a teljesítmény javulására. Némiképpen spekuláció, hiszen közvetlenül nem igazolható, de az a tény, hogy a gyermek egyre több és többféle kommunikációs helyzetben vesz részt, hozzájárul a nyelvi ingerek mind jobb feldolgozásához, és ez tükröződik egy dichotikus feladatban is. A kommunikációs helyzetek növekvő száma és sokfélesége hatással van a szókincs bővülésére, a grammatikai szerkezetek fejlődésére, a percepció működése pontosabbá válására (stb.). A gyermek egyre gyakrabban kerül olyan helyzetbe, hogy többféle nyelvi inger éri egyidejűleg (ún. kóktélpárti hatás), és egyre biztosabban lesz képes figyelni a hozzá intézett, illetve a számára fontos információkra és azok feldolgozására.

A figyelem fejlődését és hatását a dichotikus tesztek eredményeire több kutatás is magyarázó faktorként határozza meg (pl. Moncrieff 2011).

Az életkor statisztikailag is igazolt hatást gyakorolt a dichotikusan közvetített szavak helyes észlelésére. Az első hipotézisünk tehát igazolódott. A 3 évesek és a 10 évesek teljesítménye 43% különbséget mutatott (az idősebbek javára). A fejlődés a fiatalabb gyermekeknél nagyobb mértékű, mint az idősebbeknél, de végig fokozatos. Noha a helyesen visszamondott szavak száma kismértékben növekszik az életkor függvényében, azt látjuk, hogy a bal fülbe érkezőkre adott helyes válaszok növekedése valamivel kifejezettebb, mint a jobb fülben hallottak ismétlése. Ez arra utal, hogy a kétféle nyelvi inger integrációjának, illetve koordinálásának fejlődése nagyobb mértékben a bal fülbe közvetítettek pontosabb feldolgozása. Adataink hasonlóak a szakirodalomban közöltekkel (pl. Hugdahl 1999; Moncrieff 2011). A jobb fülben hallott szavak helyes ismétlése a 7–9 éves korcsoportokban azonos, a 10 éveseknél egyetlen szóval jobb. A szavak bal agyféltekei feldolgozása tehát relatíve gyorsan stabilizálódik.

A helyes válaszok minimum- és maximumértékei jellegzetesen különböznek a két fülben hallott szavak helyes visszamondásában. A jobb fülben hallottak száma már óvodáskorban lényegesen meghaladja a bal fülbe közvetítettek ismétlését. Ezek az adatok is a jobb fülből továbbított nyelvi ingerek szignifikánsan pontosabb feldolgozását igazolják.

A bevezetésben említettük, hogy a szakirodalomban közölt eredmények összevetése a módszertani (és számos egyéb) nehézség miatt csaknem lehetetlen. Ennek tudatában mégis összehasonlítottuk adatainkat egy olyan kutatásával, amelyben a vizsgált gyermekek életkora majdnem megegyezett a miénkkel, bár a dichotikus teszt egy szótagú számokat tartalmazott (Rosenberg 2011). A százalékbán megadott adatok igen nagy hasonlóságot mutattak, mind a bal, mind a jobb fül esetében. A 6, illetve a 8, 9 és 10 évesek helyes visszamondásai majdnem azonosak a két kísérletben, mind a két fülbe érkezett nyelvi ingereket tekintve. Az 5 éveseknél kisebb mértékben, a 7 éveseknél nagyobb mértékben tapasztaltunk eltéréseket a két kísérlet adataiban, főként a bal fülbe érkezett szavak esetén. Ezek az eredmények utalhatnak arra, hogy az egy szótagos számok és a két szótagból álló szavak észlelése hasonló a vizsgált populációban.

Kísérletünk részvevői döntő többségben jobbfül-fölényt mutattak, de minden korcsoportban akadtak olyanok, akikre a balfül-fölény volt jellemző, és előfordult a nem kimutatható fülfölény is. Utóbbi elenyésző mértékben. Ennek megfelelően az erre vonatkozó hipotézisünk igazolást nyert. Eredményeink igen hasonlóak az adott életkorra a szakirodalomban közölt adatokkal. Megállapítottuk, hogy a számítási módszertől függetlenül nem volt lényeges különbség a fülpreferencia alakulásában. Jellemző, hogy 8 éves kortól csökken a REA előfordulása, valamilyen mértékben növekszik a LEA előfordulása (a fiatalabbaknál tapasztaltakhoz képest) és a nem kimutatható preferencia

is. Úgy gondoljuk, hogy ez talán módszertani okokkal magyarázható; a dichotikus teszt szavainak észlelése az életkor előrehaladtával egyre könnyebb lesz. Ez eredményezi a bal fülbe közvetítettek jobb felismerését. A két fülbe közvetített nyelvi ingerek kiegyenlítettebb észlelése a beszédfeldolgozás fejlődést szemlélteti.

A 9 évesek esetében a jelentősebb mértékű LEA-előfordulásra (több mint 30%) az egyszerű kivonásos számítás esetében nem tudunk magyarázatot adni. Megjegyezzük ugyanakkor, hogy a 9 évesek beszédpercepciósi jellemzői több kutatás szerint is eltérnek a fiatalabb és az idősebb gyermekeiktől (pl. Gósy–Horváth 2006). A négyes szópárok figyelembevétele alapján nem csak a 9, de a 10 évesek esetében is csökken a REA, és növekszik a LEA és a NEA előfordulása. Ez felveti az agyféltekék, illetve a köztük lévő kapcsolat strukturális módosulásának lehetőségét mint befolyásoló tényezőt (pl. Westerhausen et al. 2011).

Angol anyanyelvű, fiatal felnőttekkel végzett CV szótagokat tartalmazó dichotikus tesztben az LI értékei 2,9 és 30,4 közöttieknek adódtak (Penna et al. 2007). Anyagunkban a lateralizációs index értékei 4,3 és 21 közöttiek, ennek alapján a gyermekek eredményei homogénebbnek tűnnek, bár a nyelvi inger jelentős különbözősége (az életkoron túl) nem hagyható figyelmen kívül.

Az LI értékeit tekintve a megállapított fülfölény, illetve agyfélteke-dominancia előfordulásának jellemzői az előzőektől lényegesen nem tértek el, igazolva ezzel azt a tényt, hogy az adott fülben helyesen visszamondott szavak száma mindenkor meghatározó. Az index figyelembevételével annyi különbséget tapasztaltunk az egyszerű kivonásos és a négyes szópárok szerinti számítás eredményeihez képest, hogy itt már a 8 éves kortól csökken jelentősen a REA előfordulása. A 9 évesek esetében az LI index kiugró LEA-előfordulást igazolt.

Ugyancsak szavakkal végzett dichotikus kísérletében Moncrieff (2011) az 5–7 éveseknél közel 60%-ban, a 8–10 éveseknél mintegy 76%-ban talált jobbfül-fölényt. Kutatásában a balfül-fölény a fiatalabbaknál mintegy 30%-ban, az idősebeknél valamivel több, mint 20%-ban fordult elő. A saját adataink – ha ugyanilyen módon összevonjuk a korcsoportokat – azt mutatják, hogy az 5–7 éveseknél a jobbfül-fölény megjelenése mintegy 87% körüli, a LEA pedig mintegy 6%-ban jelent meg. A 8–10 évesek csoportját összevonva, a REA előfordulása 67%-ban, a LEA több mint 20%-ban volt tapasztalható náluk. Az eltéréseket elsősorban a módszertani különbségek eredményezhetik. Az angol anyanyelvű gyermekek egy szótagból álló szavakat, a mieink két szótagosakat hallottak. Saját vizsgálatunkban az 5–7 évesek REA-előfordulása nagyobb mértékű volt (a különbség mintegy 27%). Ez a különbség az iskolások esetében jelentősen csökken, mindössze 9%, itt azonban a jobbfül-fölény a magyar gyermekeknél volt kisebb arányú.

Hipotézisünk szerint nem vártunk eltérést a lányok és a fiúk teljesítménye között, feltevésünk igazolódott. Megjegyezzük ugyanakkor, hogy a nemek között (használt) dichotikus vizsgálatában Moncrieff (2011) különbségeket tapasztalt, amelyek azonban az életkor szerint változtak. Adatai szerint az 5-7 év közötti lányok jobb eredményeket értek el, mint a fiúk. A 8-10 éveseknél sem a balfél-fül, sem a jobbfél-fül tekintetében nem volt különbség a lányok és a fiúk között. 11-12 éves korban a fiúknál nagyobb mértékű balfél-fül és kisebb mértékű jobbfél-fül tapasztalt a lányoknál adatolt előfordulásokhoz képest.

A jelen kutatás keresztmetszeti vizsgálat eredményeit ismertette. Ismételtén megállapíthatjuk, hogy a dichotikus vizsgálati eljárás egyszerű és relatíve gyors módszer, amely már fiatal óvodáskortól alkalmazható. Ép hallás esetén a visszamondások alapján megállapítható a fél-fül, és ennek alapján valószínűsíthető az agyfélteke-dominancia is. A jobb és bal félbe hallott szavak észlelésének, illetve ismétlésének mennyisége utal a figyelem és a memória-működés tényezőire is (a figyelem kognitív kontrolljáról lásd Sætrevik 2012-es tanulmányát). Az életkor-specifikus normák meghatározása nagyon fontos lenne, mert a dichotikus tesztek eredményei hozzájárulnának az atipikus nyelvfélldés differenciál-diagnosztikájához, és további klinikai célokra is felhasználhatók lennének.

### Irodalom

- Andrade de, Adriana Neves – Gil, Daniela – Martinelli Iorio, Maria Cecilia 2015. Benchmarks for the Dichotic Sentence Identification test in Brazilian Portuguese for ear and age. *Brazilian Journal of Otorhinolaryngology* 81/5. 459–465.
- Asbjørnsen, Arve E. – Helland, Turid 2006. Dichotic listening performance predicts language comprehension. *Laterality* 11. 251–262.
- Baddeley, Alan D. 2001. *Az emberi emlékezet*. Osiris Kiadó, Budapest.
- Best, Catherine T. 1984. Discovering messages in the medium (Speech perception and the prelinguistic infant). In Fitzgerald, Hiram E. – Lester, Barry M. – Yogman, Michael W. (eds.): *Theory and research in behavioral pediatrics*. Plenum Press, New York–London. 97–145.
- Bethmann, Anja – Tempelmann, Claus – De Bleser, Ria – Scheich, Henning – Brechmann, André 2007. Determining language laterality by fMRI and dichotic listening. *Brain Research* 1133. 145–157.
- Bever, Thomas G. 1971. The nature of cerebral dominance in speech behaviour of the child and adult. In Huxley, T. – Ingram, E. (eds.): *Language acquisition: Models and methods*. Academic Press, London. 89–103.
- Billet, Cassandra – Bellis, Teri J. 2011. The relationship between brainstem temporal processing and performance on tests of central auditory function in children with reading disorders. *Journal of Speech, Language, and Hearing Research* 54. 228–242.

- Bisazza, Angelo – Rogers, L. J. – Vallortigara, Giorgio 1998. The origins of cerebral asymmetry: a review of evidence of behavioural and brain lateralization in fishes, amphibians, and reptiles. *Neuroscience and Biobehavioral Reviews* 22. 411–426.
- Bless, Josef J. – Westerhausen, René – von Koss Torkildsen, Janne – Gudmundsen, Magne – Kompus, Kristiina – Hugdahl, Kenneth 2015. Laterality across languages: Results from a global dichotic listening study using a smartphone application. *Laterality* 20. 434–452.
- Brancucci, Alfredo – Babiloni, Claudio – Vecchio, F. – Galderisi, S. – Mucci, A. – Tecchio, F. et al. 2005. Decrease of functional coupling between left and right auditory cortices during dichotic listening: An electroencephalography study. *Neuroscience* 136. 323–332.
- Broadbent, Donald 1954. The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*. 47. 191–196.
- Brobäck, Claes-Göran 1988. *Hemispheric asymmetry and classical conditioning: an experimental study of lateralization of learning*. Reprocentralen Uppsala University, Uppsala.
- Bryden, Philip M. 1970. Laterality effects in dichotic listening: Relations with handedness and reading ability in children. *Neuropsychologia* 8. 443–450.
- Bryden, Philip M. – Allard, Fran A. 1981. Do auditory perceptual asymmetries develop? *Cortex* 17. 313–318.
- Bryden, Philip M. 1988. An introduction to the dichotic listening procedure and its relation to cerebral organization. In Hugdahl, Kenneth (ed.): *Handbook of dichotic listening: Theory, methods and research*. Chichester, UK: Wiley, 75–97.
- Carlsson, Göran – Wiegand, Gert – Stephani, Ulrich 2011. Interictal and postictal performances on dichotic listening test in children with focal epilepsy. *Brain and Cognition* 76. 310–315.
- Dawes, Piers – Bishop, Dorothy V. M. 2010. Psychometric profile of children with auditory processing disorder and children with dyslexia. *Archives of Disease in Childhood* 95. 432–436.
- Dlouha, Olga – Novak, Alexej – Vokřál, Jan 2007. Central auditory processing disorders (capd) in children with specific language impairment (SLI) - Central auditory tests, *International Journal of Pediatric Otorhinolaryngology* 71/6. 903–907.
- Doron, Karl W. – Gazzaniga, Michael S. 2008. Neuroimaging techniques offer new perspectives on callosal transfer and interhemispheric communication. *Cortex* 44. 1023–1029.
- Ettinger-Veenstra, Helene M. van – Ragnehed, Mattias – Hällgren, Mathias – Karlsson, Thomas – Landtblom Anne-Marie – Lundberg, Peter – Engström, Maria 2010. Right-hemispheric brain activation correlates to language performance. *Neuroimage* 49. 3481–3488.
- Fernandes, Myra A. – Smith, Mary Lou 2000. Comparing the fused dichotic words test and the intracarotid amobarbital procedure in children with epilepsy. *Neuropsychologia* 38. 1216–1228.
- Fernandes, Myra A. – Smith, Mary Lou – Logan, William – Crawley, Adrian – McAndrews, Mary Pat 2006. Comparing language lateralization determined by dichotic listening and fMRI activation in frontal and temporal lobes in children with epilepsy. *Brain and Language* 96. 106–114.

- Garman, Michael 1990. *Psycholinguistics*. Cambridge, Massachusetts, Cambridge University Press.
- Geschwind, Norman – Galaburda, Albert 1987. *Cerebral lateralization: Biological mechanisms, associations and pathology*. MIT Press, Cambridge.
- Gósy Mária 1996. Agyfélteke-dominancia, beszédészlelés és olvasási nehézség. In Gósy Mária (szerk.): *Gyermekkori beszédészlelési és beszédmegértési zavarok*. Nikol Kiadó, Budapest. 163–176.
- Gósy Mária 1995/2006. *GMP-diagnosztika*. Nikol Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária – Horváth Viktória 2006. Beszédfeldolgozási folyamatok összefüggései gyermekkorban. *Magyar Nyelvőr* 130/4. 470–482.
- Gósy Mária – Gyarmathy Dorottya 2011. A lateralizáció tipikus és atipikus nyelvfejlődésben. *Alkalmazott Nyelvtudomány* 9/1-2. 65–88.
- Gósy, Mária – Huntley Bahr, Ruth – Gyarmathy, Dorottya – Beke, András 2018. Dichotic listening and sentence repetition performance in children with reading difficulties. *Clinical Linguistics and Phonetics* 32. 115–133.
- Gyarmathy Éva 1998. Tanulási zavarok azonosítása és kezelése az óvodában és az iskolában. *Új Pedagógiai Szemle* 11. 68–76.
- Hagoort, Peter – Indefrey, Peter 2014. The neurobiology of language beyond single words. *Annual Review of Neuroscience* 37. 347–362.
- Hakvoort, Britt – van der Leij, Aryan – van Setten, Ellie – Maurits, Natasha – Maassen, Ben – van Zuijen, Titia 2016. Dichotic listening as an index of lateralization of speech perception in familial risk children with and without dyslexia. *Brain and Cognition* 109. 75–83.
- Halpern, Marnie E. – Güntürkün, Onur – Hopkins, William D. – Rogers, Lesley J. 2005. Lateralization of the vertebrate brain: taking the side of model systems. *The Journal of Neuroscience* 25/45. 10351–10357.
- Hámori József 1999. *Az emberi agy aszimmetriái*. Dialóg-Campus Kiadó, Budapest – Pécs.
- Helland, Turid – Asbjørnsen, Arve E. – Hushovd, Aud Ellen – Hugdahl, Kenneth 2008. Dichotic listening and school performance in dyslexia. *Dyslexia* 14. 42–53.
- Hiscock, Merrill 1988. *Behavioral asymmetries in normal children*. Molfese, Dennis L. – Segalowitz, Sidney J. (eds.): *Brain lateralization in children: Developmental implications*. Guilford Press, New York, 85–169.
- Hugdahl, Kenneth 1995. Dichotic listening: probing temporal lobe functional integrity. In Davidson, Richard J. – Hugdahl, Kenneth (eds.): *Brain asymmetry*. Cambridge, MA: MIT Press, 123–156.
- Hugdahl, Kenneth 2003. Dichotic listening in the study of auditory laterality. In Hugdahl, Kenneth – Davidson, Richard J. (eds): *The asymmetrical brain*. MIT Press, Cambridge, MA. 441–466.
- Hugdahl, Kenneth 2011. Fifty years of dichotic listening research – still going and going on. *Brain Cognition* 76. 211–213.
- Hugdahl, Kenneth – Carlsson, Göran – Uvebrant, Paul – Lundervold, Astri J. 1997. Dichotic-listening performance and intracarotid injections of amobarbital in children and adolescents. Preoperative and postoperative comparisons. *Archives of Neurology* 54. 1494–1500.

- Hugdahl, Kenneth – Brønnick, Kolbjørn – Kyllingsbæk, Søren – Law, Ian – Gade, Anders – Paulson, Olaf B. 1999. Brain activation during dichotic presentation of consonant-vowel and musical instruments stimuli: a 15O-PET study. *Neuropsychologia* 37. 431–440.
- Hugdahl, Kenneth – Heiervang, Einar – Ersland, Lars – Lundervold, Arvid – Steinmetz, Helmuth – Smievoll, Alf Inge 2003. Significant relation between MR measures of planum temporale area and dichotic processing of syllables in dyslexic children. *Neuropsychologia* 41, 666–675.
- Jerger, James – Martin, Jeffrey 2004. Hemispheric asymmetry of the right ear advantage in dichotic listening. *Hearing Research* 198. 125–136.
- Kállai János – Bende István – Karádi Kázmér – Racsmány Mihály 2008. *Bevezetés a neuropszichológiába*. Medicina Kiadó, Budapest.
- Kandel, Eric R. – Schwartz James, H. – Jessel Thomas, M. 2000. *Principles of neural science*. McGraw–Hill, New York.
- Kimura, Doreen 1961. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology* 3. 166–171.
- Krashen, Stephen D. 1973. Lateralization, language learning, and the critical period: Some new evidence. *Language Learning* 23/2. 63–74.
- Lebel, Catherine – Beaulieu, Christian 2009. Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Human Brain Mapping* 30. 3563–3573.
- Magyar Arthur 1970. A centrális hallópályák vizsgálatának újabb módszere. (Dichotikus discriminációs teszt). *Fül-Orr-Gégégyógyászat* 16. 121–124.
- Meyers, John E. – Roberts, Richard J. – Bayless, John D. – Volkert, Kurt – Evitts, Paul E. 2002. Dichotic listening: Expanded norms and clinical application. *Archives of Clinical Neuropsychology* 17. 79–90.
- Moncrieff, Deborah W. 2010. Hemispheric asymmetry in pediatric development disorders: autism, attention deficit/hyperactivity disorder, and dyslexia. In Hugdahl, Kenneth – Westerhausen, René (eds.): *The two halves of the brain*. The MIT Press, Cambridge, MA, 561–601.
- Moncrieff, Deborah W. 2011. Dichotic listening in children: Age-related changes in direction and magnitude of ear advantage. *Brain and Cognition* 76. 316–322.
- Moncrieff, Deborah – Musiek, Frank 2002. Interaural asymmetries revealed by dichotic listening tests in normal and dyslexic children. *Journal of American Academy Audiology* 13. 428–437.
- Moncrieff, Deborah W. – Wilson, Richard H. 2009. Recognition of randomly presented one-, two-, and three-pair dichotic digits by children and young adults. *Journal of the American Academy of Audiology* 20. 58–70.
- Moncrieff, Deborah – Cohen, David – Porter, Sally 2013. The psychoactive effects of psychiatric medications: The elephant in the room. *Journal of Psychoactive Drugs*, 45. 409–415.
- Musiek, Frank E. 1983. Assessment of central auditory dysfunction: the dichotic digit test revisited. *Ear and Hearing* 4/2. 79–83.
- Obrzut, John E. – Mahoney, Emery B. 2011. Use of the dichotic listening technique with learning disabilities. *Brain and Cognition* 76/2. 323–331.
- Penna, Stefania Della – Brancucci, Alfredo – Babiloni, Claudio – Franciotti, Raffaella – Pizzella, Vittorio – Rossi, Davide – Torquati, Kathia – Rossini, Paolo M. –



- Romani, Gian Luca 2007. Lateralization of dichotic speech stimuli is based on specific auditory pathway interactions: Neuromagnetic evidence. *Cerebral Cortex* 17. 2303–2311.
- Piazza, M. – Gordon, D. P. – Lehman, R. 1985. Reading ability and the development of lateralization of speech. *Language Sciences* 7. 73–84.
- Pléh Csaba 1981. Különböző szórendű mondatok értelmezése és a dichotikus hallási aszimmetriák 3-6 éves gyermekeknél. *Pszichológia* 1/3. 365–393.
- Pléh Csaba – Lukács Ágnes (szerk.) 2014. *Pszicholingvisztika*. 2014. Akadémiai Kiadó, Budapest.
- Reinhardt Melinda 2003. Változik-e az agyfélteke-dominancia kimutathatósága 5 és 10 éves kor között? *Alkalmazott Nyelvtudomány* 3/2. 91–105.
- Riès, Stéphanie K. – Dronkers, Nina F. – Knight, Robert T. 2016. Choosing words: Left hemisphere, right hemisphere, or both? Perspective on the lateralization of word retrieval. *Annals of the New York Academy of Sciences* 1369/1. 111–131.
- Rosenberg, Gail Gegg 2011. Development of local child norms for the Dichotic Digits Test. *Journal of Educational Audiology* 17. 57–63.
- Steinberg, Danny D. – Nagata, Hiroshi – Aline, David P. (eds.) 2001. *Psycholinguistics: Language, mind and world*. Longman, New York.
- Studdert-Kennedy, Michael – Shankweiler, Donald 1970. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America* 48. 579–594.
- Sætrevik, Bjørn 2012. The right ear advantage revisited: Speech lateralisation in dichotic listening using consonant-vowel and vowel-consonant syllables. *Laterality* 17. 119–127.
- Sætrevik, Bjørn – Hugdahl, Kenneth 2007. Priming inhibits the right ear advantage in dichotic listening: Implications for auditory laterality. *Neuropsychologia* 45. 282–287.
- Strouse, Anne – Wilson, Richard H. 1999. Stimulus length uncertainty with dichotic digit recognition. *Journal of the American Academy of Audiology* 10. 219–229.
- Tervaniemi, Mari – Hugdahl, Kenneth 2003. Lateralization of auditory-cortex functions. *Brain Research Review* 43. 231–246.
- Thomsen, Tormod – Specht, Karsten – Hammar, Åsa – Nyttningnes, Jarle – Ersland, Lars – Hugdahl, Kenneth 2004. Brain localization of attentional control in different age groups by combining functional and structural MRI. *NeuroImage* 22. 912–919.
- Toga, Arthur W. – Thompson, Paul M. 2003. Mapping brain asymmetry. *Nature Reviews Neuroscience* 4/1. 37–48.
- Westerhausen, René – Moosmann, Matthias – Alho, Kimmo – Medvedev, Svyatoslav – Hämäläinen, Heikki – Hugdahl, Kenneth 2009. Top-down and bottom-up interaction: manipulating the dichotic listening ear advantage. *Brain Research* 1250. 183–189.
- Willeford, Jack A. 1977. Assessing central auditory behavior in children: A test battery approach. In Keith, Robert (ed.): *Central auditory dysfunction*. Grune and Stratton, New York. 43–72.

**Dichotic listening in children: Age-specific characteristics**

The successful word recognition depends upon acquiring lexical and phonological representations during language acquisition. The intake of information through the auditory system requires an online integration of differing and potentially competing information presented to the two ears. The goal of the present study was to collect baseline developmental data on the auditory-phonetic processing of words in a dichotic listening task with the participation of 320 Hungarian-speaking children between the ages of 3 and 10. Dichotic listening techniques have been used as a sensitive non-invasive procedure to assess language lateralization. Data were scored for each participant as the percentage (and number) of correctly recalled words for the right and left ear input. In addition, three more calculations were carried to show REA, LEA and NEA with each age group. Results showed a significant increase of the correctly repeated words across ages. The difference between correct scores both in the left and right ears was also significant. As expected, more correctly recalled words were found heard in the children's right ear than in their left ear as the effect of the right ear advantage. The dichotic listening method seems to be a good way to detect the auditory-phonetic abilities of typically developing children including their interactions with memory and attention across ages.

## A BESZÉDINDULÁS KÉSÉSÉNEK HATÁSA A BESZÉDFELDOLGOZÁSRA

**Gósy Mária – Gyarmathy Dorottya**

MTA Nyelvtudományi Intézet

### **Bevezetés**

A gyermekkori beszédzavarok egyik gyakori jelensége az anyanyelv-elsajátítás megindulásának késése. Az emberi csecsemőben genetikailag kódolt a nyelvelsajátítás, annak kezdeti ideje, illetve fejlődése, a fejlődés egyes fázisai (Chomsky 1995). Az időzítésnek, az anyanyelvfejlődés egymásra épülő szakaszainak, azok lefolyásának és jellemzőinek jelentős szerepük van. Ha a tipikusnak tekintett kezdethez képest a beszéd indulása késik a gyermeknél, az számos fejlődési zavarra utalhat, amik érintik az anyanyelv-elsajátítás folyamatát. A beszéd késésének ténye más deficittek nélkül is jelentkezhet, és okozhat hosszú távú elmaradásokat, nehézségeket a későbbi életkorokban (pl. Rudolph–Leonard 2016; Zubrick et al. 2007). Az, hogy a gyermek nem az elvárt időben kezd el beszélni, jelzi a felszínen a nyelvi fejlődés átmeneti megtorpanását (Duvelleroy-Hommet et al. 1995; Kovac et al. 2001; Desmarais et al. 2008). Tovább nehezíti a késés megítélését az, hogy mit is tekintünk beszédindulásnak. Az első jelentéssel bíró szókezdeményeket? Amennyiben ezeket, akkor az elsőnek a megjelenését, avagy néhánynak a rendszeres előfordulását? Hány szónál állíthatjuk, hogy megindult a beszéd? Hogyan lehetünk biztosak az első szavak tudatos használatában? Hogyan különíthetők el az egyszerű ismétlések (imitálások) a valós produkciótól? Az említettek miatt esetleg az első rövid közlések jelentik inkább a beszédindulást? Avagy a már összefüggő beszéd? A szakirodalomban rendszerint a holofrázisok szakaszát tekintik a beszéd indulásának, vagyis az első olyan hangsorok (szókezdemények) kiejtését, amelyek a gyermek és a környezete kapcsolatában jelentést hordoznak (Gósy 2005). Természetesen ebben a szakaszban is folyamatosan bővülő „szókészletet” feltételezünk. Mindezek miatt tehát nem is olyan egyszerű a megkésett beszédindulás definiálása.

Nagyon általános megfogalmazásban megkésett beszédfejlődésről akkor beszélünk, ha a gyermek nyelvhasználata – mind a beszédprodukció, mind a beszédfeldolgozás tekintetében – elmarad az életkorában elvárt szinttől, illetve teljesítménytől, ugyanakkor sem kognitív, sem más motoros fejlődési elmaradás nem diagnosztizálható (pl. Gósy 2002a). Pusztán megfigyelés alapján a beszédprodukció elmaradott volta viszonylag jól megítélhető, a beszéd-

Gósy Mária – Gyarmathy Dorottya 2018. A beszédindulás késésének hatása a beszédfeldolgozásra. *Beszédkutatás* 2018. 186–220.

észlelés és beszédmegértés folyamatai azonban csak közvetve minősíthetők. Az utóbbiak esetében az elmaradás tehát rendkívül viszonylagos, a megítélés pedig erősen szubjektív.

A beszédindulás szempontjából a gyermeknél kétféleképpen is kialakulhat a megkésett beszédfejlődés. (i) Az egyik esetben a nyelvelsajátítás kezdete a tipikushoz képest később következik be. Ekkor a beszédindulás valamilyen mértékben elmarad a szokásostól, a gyermek 3, 4 vagy még idősebb korban kezd el beszélni (pl. Gerebenné Várbíró 1995). (ii) A másik esetben az első szavak megjelenése az elvárt időben történik ugyan, vagyis a gyermek egy- és kétéves kora között ejt szókezdemenyeket, azonban ezek mennyisége nem növekszik, az elsajátítási folyamat lelassul, és nem a tipikus anyanyelv-elsajátítás szerint zajlik.

Számos kritériumot igyekeztek meghatározni a megkésett beszédfejlődés folyamatának azonosítására (és ennek megfelelően a gyermekek diagnosztizálására) a különböző életkorokban (pl. Rice et al. 2008). Nem könnyű azonban meghatározni, hogy mely életkort vagy pontosan mit is tekintünk a nyelvelsajátítás kezdetének (pl. Perna-Loughan 2013). Az évtizedekkel ezelőtt meghatározott életkor („egyéves kor körül”) vajon érvényes-e napjainkban, avagy a gyermek számos tekintetben megváltozott verbális környezete szükségessé teszi ennek az időpontnak az újragondolását? A szakirodalom éppúgy, mint a (gyógy)pedagógia, illetve a klinikai gyakorlat többféle megközelítést, jellemzést és meghatározást alkalmaz. Általánosan elfogadott, hogy két tényezőt vesznek figyelembe, amikor a beszédindulás késését definiálják. (i) Az egyik a kétéves kori szókincs nagysága, (ii) a másik a szavak összekapcsolódásának megjelenése ebben az életkorban (Ludlow 1980; Paul-Shiffer 1991; Rice et al 2008). Vannak azonban, akik szigorúbban definiálnak, például Zubrick és munkatársai (2007). Ők a holofrázisok időszakát egyéves kor tájára teszik, a szavak összekapcsolását pedig már kétéves kor előtt várják el tipikus fejlődés esetén. Az ettől eltérő folyamatok esetén a nyelvelsajátítás késéséről beszélnek. A későbbben beszélni kezdő gyermekek között megkülönböztetnek továbbá egy olyan csoportot, akiknek a beszédindulása ugyan nem felel meg a kétéves korban elvárt kritériumoknak, de 3–5 éves korukra behozzák a kezdeti nyelvi elmaradásukat, őket az angol nyelvű szakirodalom a 'late bloomers' kifejezéssel azonosítja (pl. Thal et al. 1991). (Magyarul ez a kifejezés a 'későn érők' szókapcsolattal adható legpontosabban vissza.)

Jellemző, hogy a meghatározások, illetve a kritériumok mindig a beszéd-produkcióra vonatkoznak, aminek nyilvánvaló oka az, hogy a család, illetve a felnőttek számára a felszínen tapasztalható, azaz hallható gyermeki beszéd jelenti elsődlegesen a nyelvhasználatot. A beszédfeldolgozás, a beszédészlelés, a beszédmegértés közvetlenül nem megítélhető működések (Gósy 2005), ennélfogva kevésbé alkalmasak arra, hogy a gyermek környezete bármilyen elmaradást észrevételezzon ezekben a folyamatokban. Feltételezhető ugyanak-

kor az, hogy az anyanyelv elsajátításában tapasztalható késés, kezdeti elmaradás, avagy lassú érés éppen a beszédfeldolgozásban lenne elsőként tetten érhető, s csak ezt követően a beszédprodukciónban. Ez arra vezethető vissza, hogy a gyermek életének első heteiben már megindul a beszédpercepció fejlődése, valamivel a beszédprodukciónak fejlődést megelőzően (pl. Berko Gleason–Bernstein Ratner 1998). Ezért úgy gondoljuk, hogy annak ellenőrzése, hogy a gyermek beszédfeldolgozása megfelel-e az életkorának, rendkívül fontos. Ezeknek a folyamatoknak az objektív vizsgálata azonban módszertanilag csak idősebb életkorokban válik lehetővé.

A gyermekek – az elsajátítandó nyelvtől vagy nyelvektől függetlenül – a világon mindenütt közel ugyanabban az időben kezdenek el beszélni, ahogy Aitchison fogalmaz, mintha egy belső „biológiai óra” jelezne nekik (2003). Az nem tudható, hogy vajon ez a biológiai óra esetleg nem mindenkinek jelez, avagy vannak gyermekek, akik nem vagy csak késve reagálnak az óra jelzésére. Utóbbiak az egyébként tipikusan fejlődő, más deficitet nem mutató gyermekek, akik látszólag megértik a környezet nyelvi jelzéseit, saját beszédük kivitelezésében azonban elmaradást mutatnak. Úgy tűnik, hogy a késői anyanyelv-elsajátítást mutató gyermekek száma napjainkban sem csökken, sőt, a tapasztalatok szerint egyre gyakoribbak a kétéves korukat megközelítő, avagy két és fél éves életkorban beszélni kezdők. Amerikai adathalmazok alapján a megkésett beszédfejlődésük aránya 10% és 20% közötti; az arányok azonban nagymértékben függenek a kritériumoktól, amelyek alapján besorolják a gyermekeket (vö. Rice et al. 2008; Boyle et al. 2011).

Mind a nemzetközi, mind a hazai szakirodalomban több terminus is használatos a nem tipikus nyelvelsajátítás tényének és/vagy jellemzőinek meghatározására. A különböző terminusok kisebb-nagyobb mértékben eltérő tartalmúak (nemegyszer azonban ugyanazon a terminuson is eltérő tényeket értenek), mint például megkésett beszédfejlődés, megkésett beszédindulás, akadályozott beszédfejlődés, diszfázia, fejlődési diszfázia (pl. Wyke 1978; Crystal et al. 1989; Paul–Shiffer 1991; Duvelleroy-Hommet et al. 1995; Gerebenné 1995; Palmour 1997; Psarommatis et al. 2001; Gósy 2002a; stb.). A terminológiai bizonytalanság a jelenségnek magának a nehézkes azonosításával és többféleségével van kapcsolatban, vagyis azzal, hogy a megkésett nyelvfejlődést mutató gyermekek nagyon heterogén csoportot alkotnak. Számos szerző a megkésett nyelvfejlődés következményeit hangsúlyozza, a grammatikai elmaradásokat, avagy a tanulási nehézségeket (pl. Kraus et al. 1999). A fejlődési diszfáziát a magasabb szintű hallási folyamatok deficitjeként is kezelik, amelyek jól kimutathatók a temporális észlelés zavaraiiban (Tallal et al. 1991). Az SLI (Speech and Language Impairment), vagyis a specifikus nyelvi zavar diagnózis hátterében nagy arányban tapasztaljuk a gyermekek megkésett beszédindulását, illetve beszédfejlődését. Mindennek következménye lehet a (fejlődési) diszlexia, amelyek összefüggéseiről számos tanulmány számol be (pl. Palmour 1997; Bishop–Snowling 2004; Csépe

2007). A szerzők úgy gondolják, hogy a specifikus nyelvi zavar és a fejlődési diszlexia különböző megnyilvánulásai ugyanarra a problémára vezethetők vissza, és amelyek csak a súlyosságukban, illetve a fejlődési szintben térnek el egymástól (Bishop–Snowling 2004; Csépe 2007). Roos és Ellis Weismer tapasztalataik alapján azt találják, hogy a megkésett beszédfejlődésű gyermekek egy részét később a specifikus nyelvi zavarral diagnosztizálják (2008). Ezt igazolta Rudolph és Leonard kutatása is (2016), amelyben 150 olyan gyermeket vizsgáltak, akik egy részénél a szavak tizenöt hónapos kor után jelentek meg, míg másik részükénél a szavak összekapcsolását csak kétéves koruk után tapasztalták a szülők. A kutatás eredményei szerint a szavak későbbi megjelenése kevésbé volt rizikó tényező a későbbi nyelvfejlődés szempontjából, mint az, ha kétéves kor után jelentek meg a szókapcsolatok. Az utóbbi esetben jóval nagyobb arányban azonosítottak specifikus nyelvfejlődési zavart a gyermekeknél. A további vizsgálati eredmények szerint ugyan a genetikai és a környezeti tényezők is jelentős szerepet játszottak a zavar kialakulásában, de a meghatározó és vezető faktor a megkésett beszédindulás volt.

A (nem kívánatos) terminológiai sokszínűség jól jelzi a megkésett nyelvfejlődés jelenségének multikauzális voltát, illetőleg a különféle mintázatokat mutató természetét (Billard et al. 1996). A beszéd késése számos okra vezethető vissza, amelyek lehetnek organikusak vagy funkcionálisak, állhat a hátterükben például a hallás nem megfelelő működése, neurológiai zavar, mentális, pszichés probléma, avagy genetikai, illetve környezeti okok (pl. Billard et al. 1996; Bishop–Snowling 2004; Lyytinen et al. 2005; Preston et al. 2010; Zsoldos–Csányi 2010). Tallal és munkatársai a tipikushoz képest eltéréseket tapasztaltak a megkésett beszédfejlődésűeknél az agykéreg egyes területein, így például a parietális, a prefrontális és a temporális kéregben, és ezeket a strukturális hibákat össze tudták kapcsolni a fejlődési diszfázia profiljával (1991). A megkésett beszédfejlődés gyakran már a kezdetekben is szindróma jellegű. Az esetek jó részében ugyanakkor az ok nem vagy nem egyértelműen megállapítható.

A késői beszédindulású gyermekek egy részénél a beszédfejlődés nem mutat eltérést az időben beszélni kezdő társaikhoz képest, vagyis relatíve rövid idő elteltével (1–3 év) a nyelvi sajátosságaik az életkoruknak megfelelők lesznek (mint említettük, ők a későn érők). Náluk általában már nem lesz kimutatható a késés negatív hatása a közléseikben (pl. Ellis Weismer 2007; Paul et al. 1997; Rescorla 2002; Poll–Miller 2013). A célzott beszédészlelési terápia hiányában ugyanakkor a megkésett beszédfejlődésű gyermekek beszédfeldolgozási folyamatai hosszú éveken át nem megfelelően fejlődnek, a beszédfeldolgozás elmaradott lesz, az egyes folyamatok nem az életkornak megfelelően működnek (pl. Gósy 1998, 2002a; Zsoldos–Csányi 2010). Más szavakkal: a beszédprodukció rendeződése nem jár feltétlenül együtt a különböző beszédpercepciók részfolyamatok életkori szintű működésének kialakulásával késői beszédindulás esetén.

A szakirodalomban ismertetett kutatásokban és a mindennapi pedagógiai gyakorlatban is megkülönböztetik a beszédprodukción, illetve a beszédfeldolgozást érintő megkésett beszédfejlődést. Az adatok azt mutatják, hogy a kettő egyidejű fennállása eredményezi a legnagyobb problémát a gyermekek nyelvszajátításának folyamatára vonatkozóan: Ugyanakkor a beszédprodukción nem megfelelő fejlődése kevésbé tekinthető rizikótényezőnek az iskolai tanulási folyamatokra nézve, ha a beszédfeldolgozási mechanizmus megfelel a gyermek életkorának (Bittera–Juhász 1991; Marchman–Fernald 2013).

A megkésett beszédfejlődésű gyermekekre általánosságban az alábbiak jellemzők: a beszédfeldolgozásuk jobb, mint a beszédprodukciónjuk, a spontán narratívájuk megfelelőbb, mint a társalgásuk, a beszédük nem folyamatos, továbbá artikulációs, fonológiai, morfológiai és szintaktikai hibákat tartalmaz (pl. Gerebenné Várbió 1995; Dale et al. 2003; Lyytinen et al. 2005; Chiat–Penny 2013; Marchman–Fernald 2013). Hangsúlyozzuk, hogy nagyok ugyanakkor az individuális különbségek az adott gyermekre jellemző, a felszínen tapasztalható és a közvetlenül nem megítélhető nyelvi sajátosságok tekintetében. Sok esetben éppen ez nehezíti a késői beszédindulás negatív következményeinek időben történő felismerését.

Különböző kutatásokban vizsgálták, hogy a megkésett beszédfejlődésű gyermekek miként teljesítenek későbbi életkorokban. Egy évvel a beszéd (kései) indulása után a gyermekek 57%-a mutatott elmaradásokat, szigorúbb kritériumrendszer alkalmazásával pedig 93%-uk (Scarborough 1990). Elemezték, hogy vajon 4, illetve 5 éves korukban a késői beszédindulást mutató gyermekek milyen arányban maradnak el a kontroll csoportoktól. A kapott adatok, meglehetősen magas arányokat igazoltak, négyéves korban 71%-ot (Rescorla et al. 2000), ötéves korban 92%-ot (Ellis Weismer 2007). Szülők által kitöltött kérdőívek alapján kapott információkat összesítettek egy kutatásban (a szókincs, a nyelvtan és az absztrakt nyelv figyelembevételével), és elmaradást tapasztaltak a gyermekek 44%-ánál 3 éves korukban és 40%-uknál 4 éves korukban (Dale et al. 2003). Megkésett beszédfejlődésű, ötéves gyermekek beszédének morfológiai és szintaktikai jellemzőit vizsgálták egy kutatásban (Rescorla–Turner 2015). Noha három- és négyéves korukban ezek az angol anyanyelvű gyermekek még jelentős különbségeket mutattak a kontroll csoporthoz képest, az ötéves kori teljesítményük csak az egyes szám harmadik személyű igeragozásban mutatott szignifikáns eltérést. A mondat-tan tekintetében ugyanakkor minden területen nagy eltéréseket tapasztaltak a megkésett és a kontroll résztvevők teljesítménye között.

Több kutatás iskoláskorukban vizsgálta a (kezdetekben) megkésett beszédfejlődésű gyermekek nyelvi jellemzőit. Rescorla (2002) azt találta, hogy a 8 évesek minden elemzett nyelvi területen (szókincs, fonológia, nyelvtan, mondatisméltés) elmaradást mutattak a nem megkésett nyelvfejlődésű tanulókhöz képest. Adatai megerősítették Paul és munkatársai 1997-es eredményeit. Rice és munkatársai több kutatásban azt találták, hogy a késői beszéd-

indulás lassúbb nyelvi fejlődést eredményez, vagyis ezeknek a gyermekeknek a teljesítménye fiatalabb kortársaikéval egyezik meg (Rice et al. 2006; Rice 2007). 128 megkésett beszédfejlődésű gyermek nyelvi jellemzőit vizsgálták 7 éves korukban, korban nemben illesztett kontroll csoporttal összehasonlítva (Rice et al. 2008). Megállapították, hogy a kétévesen beszélni kezdő gyermekek szignifikánsan elmaradnak a vizsgált nyelvi teljesítményükben az időben beszélni kezdőktől. Jellemzően nagy különbségeket tapasztaltak a szintaxis terén, különösen a grammatikai időviszonyok kifejezésében. Nem igazolódott azonban különbség a lányok és a fiúk között. Az olvasási és a matematikai teljesítményben mutattak ki szignifikáns különbséget a tipikus fejlődésű és a megkésett nyelvelsajátítású gyermekeknek 10 éves korukban (Perna–Loughan 2013). Nyolcéves, gyenge nyelvi készséggel rendelkező gyermekeket és tipikus fejlődésűeket vizsgáltak egy kutatásban, majd ezt követően azonosították az anyanyelv-elsajátítási „történetüket” (Poll–Miller 2013). Azt ellenőrizték, hogy vajon kétéves korukban mekkora volt a szókincsük, illetve hogy képesek voltak-e már összekapcsolni a szavakat a közléseikben. Az eredmények szerint a nyolcévesen gyenge nyelvi teljesítményt mutató gyermekek több mint 60%-a nem volt megkésett beszédindulású. Ugyanakkor az a tény, hogy a gyermek kétévesen nem kapcsolta össze a szavakat, egyértelműen rizikó faktornak minősült a későbbi gyenge nyelvi teljesítmény szempontjából. A tanulmány szerzői arra a következtetésre jutottak, hogy az iskoláskorban tapasztalt gyenge nyelvi teljesítményhez többféle út vezethet. Ezek egyike a késői beszédindulás. Huszonnyolc megkésett beszédfejlődésű és huszonöt kontroll gyermek vett részt Rescorla (2005) egyik vizsgálatában. Valamennyien 13 évesek voltak. A megkésettnek minden elemzett területen gyengébben teljesítettek, mint a kortársaik, különösen a szókincs, a nyelvtan, a verbális emlékezet és az olvasásértés tekintetében, ugyanakkor az eredményeik a sztenderd tesztekben az életkorukban elvárt szintnek megfelelőek. A szerző úgy fogalmazott, hogy a 2 éves és 2,5 éves kor közötti lassabb nyelvelsajátítás hosszú évekkal később is negatív hatással van bizonyos nyelvi teljesítményekre.

A kutatások döntően a megkésett beszédfejlődésű gyermekek beszédprodukciós sajátosságaira fókuszáltak, jóval kevésbé a beszédfeldolgozásra. A késői beszédindulást mutató gyermekek egy részénél nem találtak (jelentős) eltérést a beszédészlelés és a beszédmegértés vizsgált folyamataiban a későbbi életkorokban (pl. Martin 1981), míg más részükénél jelentős zavarokat tapasztaltak, főként az olvasott szövegek megértésében (pl. Gósy 1998, 2002a; Lyytinen et al. 2005; Horváth 2007; Preston et al. 2010; Poll–Miller 2013; Perna–Loughan 2013; Brooks–Kempe 2014).

Jellemző nehézségeket tapasztaltak a megkésett beszédfejlődésűeknél például a mondatismétlésben (Ellis Weismer 2007; Roos–Ellis Weismer 2008), illetve iskoláskorban az olvasottak megértésében (Rescorla 2005; Poll–Miller 2013; Perna–Loughan 2013). Az értelmetlen szavak ismétlése is jóval nehe-



zebbnek bizonyult a megkésett beszédfejlődésűek számára a tipikus nyelvelsajátítást mutatókhoz viszonyítva (Conti-Ramsden et al. 2001). A kutatók az okot abban látják, hogy az értelmetlen szavak ismétlése meglehetősen összetett folyamat, amely észlelési működéseket és a fonológiai formák azonosítását egyaránt tartalmazza, továbbá szükséges hozzá a rövid idejű memória megfelelő kapacitása is (Bishop 2006). Az ilyen tesztekben nyújtott teljesítmény jól jelzi a háttérben lévő nyelvi nehézségeket (Gósy 2002a, 2002b; Gósy–Horváth 2006; Rosta–Schuchné Rumpli 2007; Horváth 2007; Chiat–Roy 2013). Longitudinális vizsgálatban elemezték a mandarin lexikális dallamminták megkülönböztetésének biztonságát olyan gyermekeknél, akik két éves koruk körül kezdtek beszélni. (A tonális nyelvekben a lexikális dallam jelentéseket különböztet meg.) A neuropszichológiai elemzéseket a gyermekek 3, 5 és 6 éves korában végezték (Chen et al. 2016). A megkésett beszédfejlődésűeket két csoportra osztották, az egyiket azok alkották, akiknél a nyelvelsajátítás késése folyamatosan fennállt, míg a másikat azok, akik a vizsgálat időpontjaiban már nem mutattak szignifikáns eltérést az alkalmazott nyelvi tesztekben a tipikus fejlődésű gyermekekhez képest (utóbbiak voltak az ún. későn érők). Az eredmények azt mutatták, hogy a vizsgált életkorokban jelentős eltéréseket tapasztaltak azoknál a gyermekeknél, akik folyamatos késést mutattak szemben azokkal, akik behozták az elmaradásukat (bár a különbségek sajátosan módosultak az életkor előrehaladtával).

Scarborough és Dobrich (1990) felvetik az 'illuzórikus javulás' fogalmát, ami akkor következik be, amikor a megkésett beszédindulást látszólag tipikus nyelvelsajátítás követi, valójában azonban iskoláskorban súlyos deficitet mutathatók ki ezeknél a gyermekeknél, amelyek elsősorban az írott nyelv megtanulását érintik. Perna és Loughan (2013) úgy fogalmaznak, hogy a késői beszédindulás mindig rizikó tényező marad a gyermek későbbi nyelvelsajátítását, illetve az írott nyelv birtokba vételét, sőt az általános tanulási folyamatokat tekintve is. Úgy gondoljuk, hogy ebben fontos tényező a beszédfeldolgozás, amelynek elmaradásairól és azok mértékéről kevesebb objektív adat áll rendelkezésre. Úgy tűnik, hogy a szakemberek pillanatnyilag nem tudják megmondani, hogy a késés milyen mértékben érinti azokat a gyermekeket, akikre 2 és 3 éves koruk között nem a tipikus nyelvfejlődés volt jellemző. Az sem egyértelmű még, hogy melyek azok a nyelvi és egyéb területek (pl. figyelemkoncentráció problémája, pszichés gondok, szorongás), amelyek jobban, és melyek azok, amelyek kevésbé érintettek. Saját tapasztalatunk azt igazolja, hogy például a beszédprodukció látszólagos életkori szintje mellett a beszédészlelési és/vagy beszédmegértési folyamatok nem életkor-specifikus működése jellemző a megkésett beszédindulás esetén (Gósy 2002a, 2002b; Rosta–Schuchné Rumpli 2007; Horváth 2007). A későbbi életkorokban a gyermek fennálló problémáit már kevésbé kapcsolják össze a késői beszédindulással, hanem egyéb diagnózist adnak, mint például a specifikus nyelvi zavar, tanulási nehézség, figyelemzavar stb. Egy utánkövetéses vizsgálat ered-

ménye szerint a késői beszédindulású gyermekek 39%-át figyelmi deficittel, illetve hiperaktív zavarral diagnosztizálták (ADHD: Attention Deficit Hyperactive Disorder, vö. Perna–Loughan 2013). Harminc megkésett beszédfejlődésű gyermeket újravizsgáltak hatéves korukban (Miniscalco et al. 2006). A gyermekek 71%-ánál valamilyen mentális zavart találtak, beleértve az autizmus spektrum zavart (5%-ban). 11%-uk mutatott figyelmi deficitet, 7%-ukat fejlődési koordinációs zavarral, 8%-ukat határeseti IQ-val, illetve enyhe tanulási zavarral és 3%-ukat olvasászavarral diagnosztizáltak. Perna és Loughan (2012) alacsonyabb intelligencia- és memóriaértékeket talált a megkésett beszédfejlődésű gyermekeknél.

Láttuk, hogy sok kutatás foglalkozik a megkésett beszédindulású gyermekek nyelvi teljesítményével későbbi életkorukban. Felfogásunkban a beszéd kezdetét az jelentette, amikor a gyermek elkezdte a hangsorokat szó funkcióban használni, vagyis megkezdődött a holofrázisok időszaka. Felmerül a kérdés, hogy vajon hogyan működik a 3 és 4 éves korukban beszélni kezdő magyar anyanyelvet elsajátító gyermekek beszédfeldolgozási mechanizmusa. Kutatásunkban ezeknek a gyermekeknek a beszédészlelési és beszédmegértési sajátosságait tanulmányoztuk hatéves korukban.

Kutatásunk célja annak megismerése, hogy a beszédindulás eltérő időpontjai milyen hatással vannak (hatással vannak-e) a beszédfeldolgozási folyamatok működésére hatéves korban, vagyis az iskolakezdést közvetlenül megelőzően. Három hipotézist fogalmaztunk meg. Feltételeztük, hogy (i) a késői beszédindulás következtében elmaradásokat fogunk tapasztalni a gyermekek hatéves kori beszédészlelési és beszédmegértési folyamataiban (a kontroll csoporthoz képest), (ii) azok a gyermekek, akik négyéves korukban kezdtek el beszélni, súlyosabb beszédfeldolgozási elmaradásokat mutatnak, mint azok, akik háromévesen kezdtek beszélni; (iii) a későn beszélni kezdők beszédfeldolgozási folyamataiban különböző mértékű elmaradásokat fogunk tapasztalni.

### **Kísérleti személyek, anyag, módszer**

#### **Kísérleti csoportok**

Kutatásunkban összesen 60 gyermek vett részt, a vizsgálatkor mindannyian 6 évesek voltak ( $\pm 3$  hónap). A nemek eloszlása mindhárom csoportban 75%–25% volt a fiúk javára, tehát minden csoportba 15 fiú és 5 lány tartozott. Ez az arány jellemzőnek mondható a megkésett beszédfejlődésű gyermekek esetében. A résztvevők három csoportot alkottak, két csoportban megkésett beszédfejlődésű gyermekek voltak (összesen negyvenen), a harmadikban pedig korban, nemben illesztett, tipikus anyanyelvfejlődést mutató kontrollok (összesen húsz gyermek). A megkésett beszédfejlődésűek egyik csoportjába azokat a gyermekeket soroltuk, akik 3 évesen, a másik csoportba pedig azokat, akik 4 éves korukban kezdtek el beszélni.

A beszédindulás minden gyermek esetében tehát a szavak (szókezdemények) ejtését jelentette, az életkort pedig a szülők közölték. A szavak összekapcsolásának időpontja a gyermekek többségénél a szülők bizonytalan válaszai miatt nem volt egyértelműen azonosítható. (Néhány szülő a 3, illetve 4 éves korban megkezdett logopédiai fejlesztést követő hónapokhoz kötötte a szavak összekapcsolásának időpontját.) A nem tipikus anyanyelv-elsajátítási folyamatot mutató gyermekek valamennyien ép hallók voltak (részben klinikai audiométerrel, részben a GOH-eljárással mérve). A WISC-IV intelligenciateszt értékei minden gyermeknél az ép övezetbe estek. A 40 megkésett beszédfejlődésű gyermek közül huszonnyolcnak tiszta volt az artikulációja, 12 gyermeknél a sziszegő-susogó mássalhangzók és közülük 6 gyermeknél még a pergőhang képzése volt hibás. Egyik gyermeknél sem merült fel az agrammatizmus; beszédük nyelvtani tekintetben a logopédus véleménye szerint megfelelt az életkorukban elvárt szintnek. Valamennyi gyermek 1, illetve 2 éven át járt logopédushoz (öt gyermekkel a vizsgálat időpontjában is foglalkozott szakember). A kontrollcsoportot ép halló és ép értelmi képességű, tipikus fejlődésű gyermekek alkották. Valamennyien egy- és másfél éves koruk között ejtették az első szavakat, azt követően pedig az anyanyelv-elsajátításuk folyamatos és problémamentes volt.

#### **A kutatás módszeréről**

A gyermekek beszédészlelési és beszédmegértési folyamatait a GMP-diagnosztika (Gósy 1995/2006) 20 tesztje közül nyolcnak az alkalmazásával vizsgáltuk. A GMP-diagnosztika sztenderdizált eljárás, amelynek segítségével a beszédészlelési és a beszédmegértési folyamat működésének szemszögből vizsgálható a magyar gyermekek anyanyelv-elsajátítása. A sztenderd értékek révén megítélhető, hogy a gyermek teljesítménye életkor-specifikus-e vagy nem.

A jelen kutatásban felhasznált nyolc teszt közül hat a beszédészlelési folyamatokat (GMP2, 4, 5, 10, 17, 18), kettő pedig a beszédmegértést (GMP12, 16) vizsgálta. A GMP2 az akusztikai-fonetikai észlelés működését, a GMP4 a fonetikai észlelést, illetve az akusztikai kulcsok felismerését, a GMP5 a morfo-fonológiai észlelést, a GMP10 a szeriális észlelést, a GMP17 pedig a beszédhang-differenciálást tesztelte. A GMP18 a transzformációs észlelésre fókuszált. A beszédmegértést ellenőrző tesztek közül a GMP12 a szövegértésről, a GMP16 a mondatértésről nyújtott információt. A következőkben a tesztek és az eljárás részletes ismertetése következik.

#### **Az alkalmazott tesztek és a tesztelési eljárás**

Három tesztben (GMP2, 4, 5) a gyermekek feladata 10-10 mondat ismétlése volt, amelyek – több szempontból is – kiváló diagnosztikai lehetőséget nyújtanak a különféle nyelvelsajátítási zavarok kimutatásához (pl. Rescorla 2002; Roos–Ellis Weismer 2008; Seeff–Gabriel et al. 2010). A mondatismét-

léses feladatokat széles körben használják a klinikai gyakorlatban, és hasznosnak bizonyultak a nyelvfeldolgozási gyengeség kimutatására (Klem et al. 2015). A GMP2 a beszédészlelés akusztikai-fonetikai szintjeinek felismerésére irányul. A gyermekeknek 10 db, fehér zajjal elfedett (jel/zaj viszony 4 dB) mondatot kell megismételniük (egyszeri hallás alapján, azonnal). A mondatokat férfi beszélővel rögzítették. A tíz mondat helyes ismétlésének sztenderd értéke 6 éves korban 90%. Példák: *Rakjátok össze a játékokat!*, avagy *A repülőgép most szállt le.*

A GMP4-ben 10 db szűk frekvenciás mondat azonosítása történik (mintegy 500 Hz-es frekvenciatartományú, 2200 Hz és 2700 Hz közötti, 36 dB/oktáv meredekségű szűrő alkalmazásával létrehozott mondatok). A teszt eredménye az akusztikai kulcsok felismerésére, illetve a fonetikai feldolgozásban való felhasználásuk sajátosságaira világít rá. Példák: *A tejfölt a macska itta meg.*, *Dobd a papírt a szemétkosárba!* Az elvárt teljesítmény 6 éves kortól 100%. A gyermek feladata a mondatok egyszeri hallás alapján történő azonnali ismétlése.

A GMP5-ös tesztben 10 db gyorsított mondat azonosítása a feladat (ugyancsak egyszeri hallás alapján, azonnali ismétléssel). A mondatokat, amelyeket átlagos beszédtempójú férfi beszélő ejt, 20%-kal gyorsították fel műszeresen. Itt a jelentés és az asszociációs szintek részleges kizárásával mérhető fel a morfo-fonológiai teljesítmény. 6 évesek esetében 90% az elvárt teljesítmény. Példák: *Az irigység rossz tulajdonság.*, *A katonák felesküdték a zászlóra.*

Az életkor-specifikus sorozatészlelés helyes működése, valamint a megfelelő beszédhang-differenciálás elengedhetetlenek a mentális lexikon növekedéséhez, valamint a nyelv írott formájának megtanulásához. A szeriális észlelés tesztjében (GMP10) értelmetlen hangsorok visszamondatásával vizsgáljuk a sorozatészlelés működését. Példák: *galalajka*, *feréndekek*, *vucsityó*. Hatéves kortól 100%-os teljesítményt várunk el. Az értelmetlen hangsorok pontos felismeréséhez a megfelelő kapacitású rövid idejű memória is szükséges (pl. Conti-Ramsden et al. 2001; Bishop 2006; Chiat–Roy 2013).

A GMP17-es teszt a beszédhang-megkülönböztetés működésének vizsgálatára irányul. A vizsgáló olyan értelmetlen, egy vagy két szótagú hangsorokat olvas fel a gyermeknek, amelyek vagy azonosak (pl. *hem* vs. *hem*), vagy csak egy fonológiai jegyben különböznek egymástól (pl. *nőcs* vs. *nőcs*, *aszá* vs. *asá*, *gev* vs. *bev*, *ippi* vs. *ipi*). A gyermeknek kizárólag az észlelésére hagyatkozva kell eldöntenie, hogy az adott hangsorok azonosak vagy különbözők voltak-e. Hatéves korban csupán a mássalhangzók időtartamában lehet bizonytalan a gyermek; ennek megfelelően az elvárt érték 19 teszt-pár helyes azonosítása a 23-ból (a sztenderd érték tehát 82,6%).

A transzformációs észlelés (GMP18) vizsgálata során arra kapunk választ, hogy képes lesz-e a gyermek az anyanyelvi beszédhangokat (később, az olvasástanuláskor) a betűknek megfeleltetni. Alkalmazza-e az elhangzásnak meg-

felelő beszédhang-differenciálást és sorozatészlelést egy vizuális/taktilis transzformáció során, amely elengedhetetlen az írás-olvasás megtanulásához. Hatéves kortól 75%-os teljesítmény az elvárt szint.

A szövegértést (GMP12) egy elhangzott meséhez kapcsolódó kérdésekre adott válaszok alapján, a mondatértést (GMP16) pedig az elhangzottak színes rajzokhoz történő társítása révén ítéljük meg. A szövegértési teszt a szemantikai szerkezetek értelmezésének, a részletek azonosításának, az ok-okozati viszony felismerésének, illetve az asszociációs szint működésének vizsgálatára irányul. A hallott mese mintegy 2 perc időtartamú, ennek elhangzása után a gyermeknek 10 kérdésre kell adekvát választ adnia. Hatéves korban 70-80%-os eredmény az elvárható.

A mondatértés vizsgálata (GMP16) az adott életkorban szükséges szemantikai sajátosságok és a grammatikai struktúrák feldolgozásába enged betekintést. A teszt eszköze 10 db rajzpár; az egyes párok vizuálisan mindig csak egyetlen dologban különböznek oly módon, hogy az adott képpárhoz tartozó tesztmondat mindig csak az egyik képre igaz (pl. *A nyuszt kergeti az oroszlán.* tesztmondathoz tartozó képek egyikén a nyuszt kergeti az oroszlán, a másikon a nyuszi kergeti az oroszlánt; avagy az *A kislánynak oda kell adnia a könyvet a kisfiúnak.* tesztmondathoz tartozó egyik képen a kislány nyújtja a könyvet, a másikon pedig a kisfiú, vö. 1. ábra). A gyermeknek mindig az adott tesztmondat tartalmának megfelelő képet kell megmutatnia. Hatéves kortól 90-100%-os teljesítmény várható el.



1. ábra

Példa a mondatmegértési teszthez használt képpárra (az elhangzó tesztmondat: *A kislánynak oda kell adnia a könyvet a kisfiúnak.*)

A tesztek felvétele minden gyermek esetében délelőtt zajlott, ismert környezetben, a tesztelés teljes időtartama 20 perc és 30 perc közötti volt.

Elemeztük az egyes tesztekben elért teljesítményt csoportszinten, illetve a három csoport összevetésében, továbbá a sztenderd értékek mentén is. Az adatok statisztikai vizsgálatát az SPSS 20-as szoftverrel végeztük nem-paraméteres Chi-négyzet goodness-of-fit teszttel (az adatok véletlenszerű vagy nem véletlenszerű eloszlására nézve; az egyes teszteknek csak a nem véletlenszerű eloszlást közöltük). Általános lineáris kevert modellt alkalmaztunk (GLMM), amelyben a függő változó az egyes tesztekben elért helyes válaszok aránya, a független változók pedig a beszédindulás ideje, a teszt típusa, illetve a nem (lány vagy fiú) voltak. A nemek szerinti elemzést annak tudatában végeztük el, hogy a fiúkat számszerűen jóval nagyobb mértékben érintette a beszédindulás késése, mint a lányokat, ezért adataik összehasonlítását tájékoztató jellegűnek tekintjük.

### Eredmények

A kutatásba bevont nyolc teszt eredményeire épített általános lineáris kevert modell alkalmazásának eredményeként a statisztikai elemzés kimutatta, hogy a három vizsgált csoport között az összes teszt eredményét együtt elemezve szignifikáns különbség van ( $F(2, 446) = 139,795$ ;  $p = 0,001$ ). A további elemzéseket a következők szerint végeztük. Minden teszt esetében páronként is összevetettük a csoportok teljesítményét, elemeztük az egyéni különbségeket, és összehasonlítottuk a csoportok eredményeit az életkorukban elvárt, sztenderd értékekkel. Elemeztük a lányok és a fiúk teljesítményét minden csoportban és minden tesztben. A következőkben részletesen elemezzük az egyes tesztek eredményeit a teszt (a diagnosztikában szereplő) sorszáma szerint.

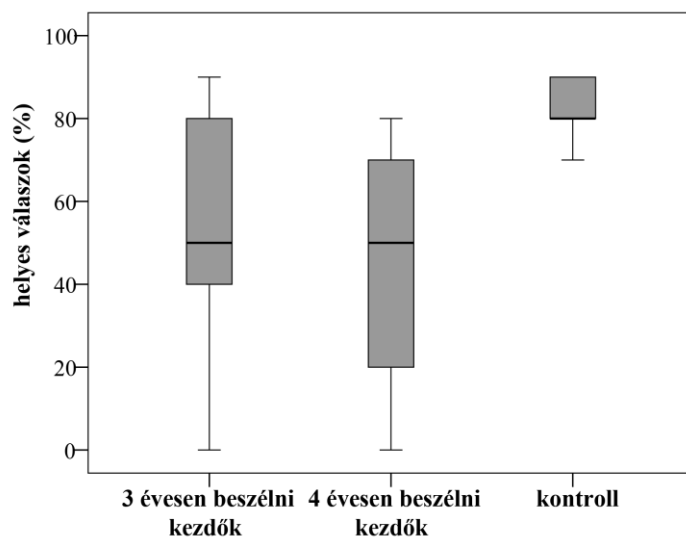
### GMP2

A beszédészlelés akusztikai-fonetikai szintjét vizsgáló tesztben a csoportok egyikének átlaga sem érte el az életkorban elvárt 90%-ot. A legjobban – várhatóan – a kontrollcsoportba tartozó gyermekek teljesítettek (átlag: 81,5%, átlagos eltérés: 7,5%), de közülük is a 20-ból mindössze 7 gyermek (a csoport 35%-a) teljesítménye volt 90%. További 9 gyermeknél 80%-os (a csoport 45%-a), míg 4 gyermeknél (a csoport 20%-a) 70%-os teljesítményt tapasztaltunk. A lányok (átlag: 86%, átlagos eltérés: 5,5%) 6 százalékponttal teljesítettek jobban a fiúknál (átlag: 80%, átlagos eltérés: 7,6%).

Azoknak a gyermekeknek a teljesítménye, akik 3 éves korukban kezdtek el beszélni, átlagosan 57%-os volt (átlagos eltérés: 24,6%). Ebben a csoportban 4 gyermek érte el az életkorilag elvárt 90%-ot, kettő a 80%-ot, három a 60%-ot, öt az 50%-ot, négy a 40%-ot, és egy-egy gyermek nyújtott 20, illetőleg 0%-os teljesítményt. A fiúk (átlag: 53%, átlagos eltérés: 18,4%) és a lányok (átlag: 66%, átlagos eltérés: 39,1%) átlagos teljesítménye mintegy 13 százalékpontos eltérést mutat. Érdeemes megjegyezni, hogy a legjobb és a leggyengébb teljesítmény is a lányoknál volt adatolható. A leggyengébbeknek a

tesztben azok a gyermekek mutatkoztak, akiknek a beszédindulása négyéves korra tehető, az ő eredményük átlagosan 43%-nak adódott (átlagos eltérés: 28,1%). Ez utóbbi csoportban egyetlen gyermeknek sem sikerült elérnie az elvárt szintet. A legjobb, 80%-os teljesítmény mindössze két gyermeknél volt adatolható, öten nyújtottak 70%-os, hárman-hárman 50%-os, 30%-os, illetve 0%-os teljesítményt, egy-egy gyermek 60%-osat, illetve 10%-osat, míg 20%-ot két gyermek ért el. A fiúk teljesítménye (átlag: 39%, átlagos eltérés: 26,6%) ebben a csoportban is jóval alacsonyabb volt, mint a lányoké (átlag: 54%, átlagos eltérés: 32,9%). A nemek között azonban egyik csoportban sem volt matematikailag igazolható különbség.

A három csoport eredményei szignifikánsan különböznek egymástól ( $F(2, 446) = 17,530$ ;  $p = 0,001$ ). A páronkénti összehasonlítás pedig megerősítette, hogy a 3 és a 4 évesen beszélni kezdők csoportja között ( $t = 2,189$ ;  $p = 0,029$ ), a 3 évesen beszélni kezdők és a kontroll csoport ( $t = 3,670$ ;  $p = 0,001$ ), valamint a 4 évesen beszélni kezdők és a kontroll csoport között is ( $t = 5,859$ ;  $p = 0,001$ ) szignifikáns különbség áll fenn (2. ábra).



2. ábra

A mondatazonosítás zajban (GMP2) tesztjének eredményei a három csoportban (medián és szóródás)

Az egyes csoportok eredményeit (a modellen belül a páronkénti összehasonlítással) összevetettük az életkorban elvárt sztenderd értékekkel. A statisztikai elemzés szerint csak a kontrollcsoport teljesítménye nem mutat szignifi-

káns elmaradást az elvárt szinttől, ellentétben a 3 évesen ( $t = 5,529$ ;  $p = 0,001$ ), illetve a négyévesen ( $t = 7,999$ ;  $p = 0,001$ ) beszélni kezdő gyermekekkel.

#### GMP4

A fonetikai észlelést és az akusztikai kulcsok felismerését vizsgáló tesztben szintén a kontrollcsoportba tartozó gyermekek teljesítettek a legjobban, a csoportátlaguk 96% volt (átlagos eltérés: 5%), amely csak 4 százalékponttal marad el az életkorilag elvárt szinttől. A csoport 60%-a nyújtott 100%-os teljesítményt, míg a 40%-uk 90%-osat. A lányok (átlag: 98%, átlagos eltérés: 4,5%) és a fiúk (átlag: 95,3%, átlagos eltérés: 5,2%) eredménye csupán alig 3 százalékpontos eltérést mutatott.

A 3 évesen beszélni kezdők csoportszinten 67%-os eredményt (átlagos eltérés: 29,8%) értek el. A 6 éves korban elvárt hibátlan teljesítményt a csoportból két gyermek érte el csupán; ugyanennyien produkáltak 60%-os, 30%-os, illetve 0%-os eredményt. Négy-négy gyermek pedig 90%-os, 80%-os, illetőleg 70%-os teljesítményt ért el. A lányok (átlag: 70%, átlagos eltérés: 41,2%) átlagosan négy százalékponttal jobb eredményt értek el, mint a fiúk (átlag: 66%, átlagos eltérés: 26,7%). A leggyengébben itt is azok a gyermekek teljesítettek, akiknek a beszédindulása 4 éves korra tehető; ők csoportszinten az elhangzott mondatok alig több, mint felét (átlag: 51% átlagos eltérés: 28,8%) tudták helyesen megismételni. Nem akadt hibátlan megoldás; a legjobb, 80%-os eredményt mindössze négyen érték el. Hatan 70%-ot, ketten-ketten 60%-ot, 10%-ot és 0%-ot, míg egy-egy gyermek 50%-ot, 40%-ot, 30%-ot, illetve 20%-ot ismételt helyesen. A fiúk (átlag: 47,3%, átlagos eltérés: 28,7%) és a lányok (átlag: 62%, átlagos eltérés: 29,5%) közötti különbség nem volt szignifikáns.

A csoportok átlagteljesítménye matematikailag igazolhatóan különbözik egymástól ( $F(2, 446) = 24,080$ ;  $p = 0,001$ ). A páronkénti összehasonlítás minden esetben szignifikáns különbséget igazolt. A 3 évesen és a 4 évesen beszélni kezdők között:  $t = 2,577$ ;  $p = 0,010$ ; a 3 évesen beszélni kezdők és a kontrollcsoport között:  $t = 4,291$ ;  $p = 0,001$ ; illetve a 4 évesen beszélni kezdők és a kontrollcsoport között:  $t = 6,869$ ;  $p = 0,001$  (3. ábra). A három csoport eredményeit (matematikailag) összevetettük az elvárt (sztenderd) eredményekkel, ami kimutatta, hogy a kontrollcsoport marad el a legkevésbé a kívánt szinttől, a megkésett beszédindulású gyermekek mindkét csoportja azonban jelentős elmaradást mutat (3 évesen beszélni kezdők:  $t = 5,442$ ;  $p = 0,001$ ; 4 évesen beszélni kezdők:  $t = 8,350$ ;  $p = 0,001$ ).

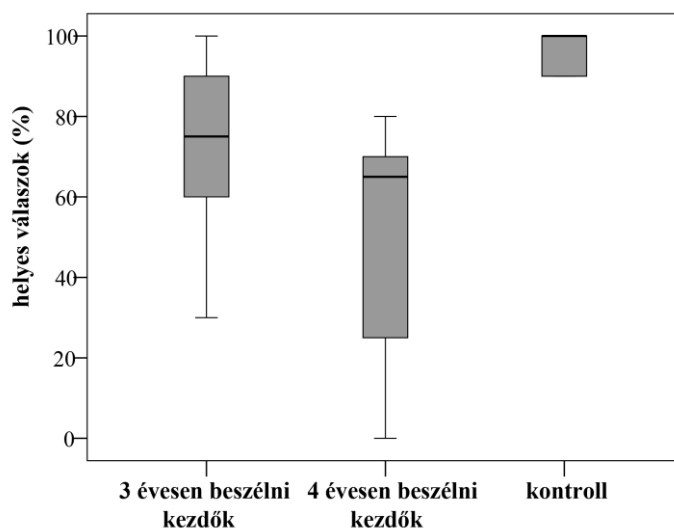
#### GMP5

A beszédészlelési teljesítményt a jelentés és a szóasszociációs szintek részleges kizárásával vizsgálja ez a teszt. Itt adatoltuk az egyik leggyengébb teljesítményt mindhárom csoportban. Noha – elvárásainknak megfelelően – itt



is a kontrollescsoport eredménye volt a legjobb, a hatéves korban elvárt 90%-hoz képest mindössze átlagosan 65,5%-ot (átlagos eltérés: 9,4%) érték el a gyermekek csoportszinten. A legjobban teljesítő óvodások (3 fő) a mondatok 80%-át tudták helyesen visszaadni; nyolcan 70%-ot, hatan 60%-ot, hárman pedig 50%-ot értek el. A lányok (átlag: 66%) és a fiúk eredménye (átlag: 65,3%) nem különbözött egymástól.

A három-, illetve a négyévesen beszélni kezdők csoportja egyaránt 27%-os átlagértéket (átlagos eltérés a 3 évesen beszélni kezdőknél: 20,6%, átlagos eltérés a 4 évesen beszélni kezdőknél: 29,6%) produkált. A 3 évesen beszélni kezdők csoportjában a legjobb eredmény 80% volt, amit egyetlen gyermek ért el, a leggyengébb pedig 0%, amit két esetben adatoltunk. Három-három gyermek nyújtott 50%-os, illetve 40%-os, egy 30%-os, míg öt-öt gyermek 20%-os, illetve 10%-os teljesítményt. A lányok 10 százalékponttal (átlag: 34%, átlagos eltérés: 32,1%) voltak jobbak a fiúknál (átlag: 24%, átlagos eltérés: 16%). A nemek tekintetében nem igazolódott szignifikáns különbség.



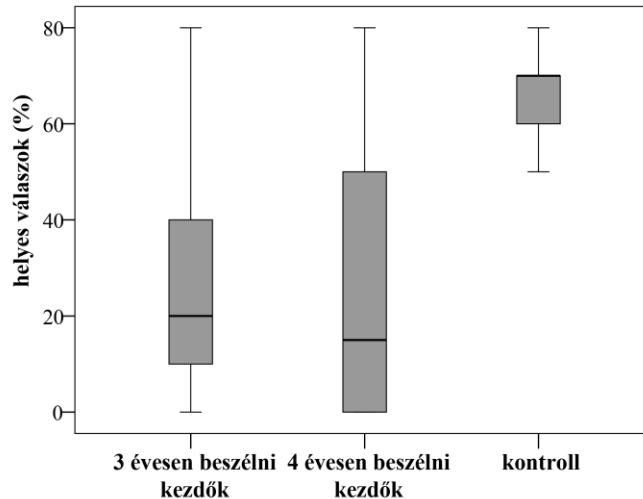
3. ábra

A szűk frekvenciás mondatok helyes ismétlésének (GMP4) eredményei a három csoportban (medián és szóródás)

A négyévesen beszélni kezdőknél szintén 80% volt a legjobb teljesítmény, amelyet három gyermek is elért, de náluk a csoport több mint negyede (6 gyermek) produkált 0%-ot. Ketten-ketten a mondatok 60%-át, illetve 40%-át, hárman a 20%-át, négyen pedig a 10%-át tudták visszamondani. A lányok

teljesítménye (átlag: 32%, átlagos eltérés: 35,6%) ebben a csoportban is jobb volt, mint a fiúké (átlag: 25,3%, átlagos eltérés: 28,5%). A nemek között egyik csoportban sem volt szignifikáns különbség megállapítható.

A három csoport átlagteljesítménye szignifikánsan különbözik egymástól ( $F(2, 446) = 22,830$ ;  $p = 0,001$ ), de a csoportok páronkénti összehasonlítása csak a három évesen beszélni kezdők és a kontroll csoport ( $t = 5,845$ ;  $p = 0,001$ ), illetve a négyévesen beszélni kezdők és a kontroll csoport között ( $t = 5,859$ ;  $p = 0,001$ ) igazolt matematikai különbséget (4. ábra). A kapott adatokat összevetettük a sztenderd értékekkel, ami azt igazolta, hogy mind a kontroll csoportba tartozó ( $t = 4,192$ ;  $p = 0,001$ ), mind a három- ( $t = 10,787$ ;  $p = 0,001$ ), mind pedig a négyévesen beszélni kezdők eredményei ( $t = 10,803$ ;  $p = 0,001$ ) szignifikánsan különböznek az ebben az életkorban elvárt teljesítménytől.



4. ábra

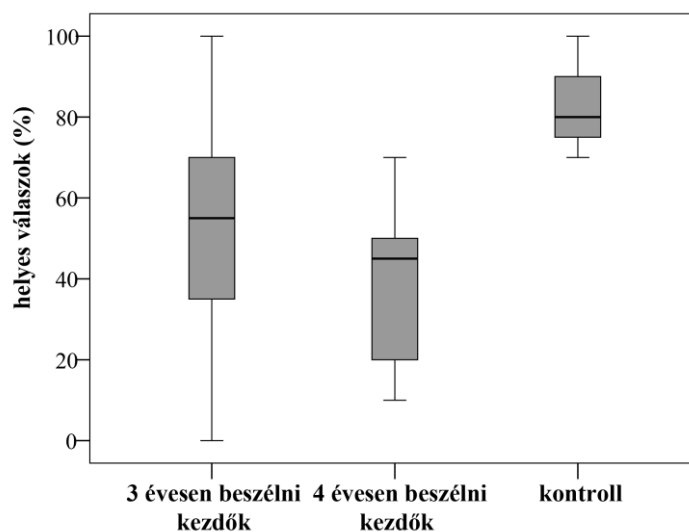
A morfo-fonológiai komplex mondatok ismétlésének (GMP5-ös teszt) eredményei a három csoportban (medián és szóródás)

### GMP10

A sorozatészlelés működését felmérő tesztben hatéves kortól 100% az elvárt teljesítmény, amelyet a legjobban a kontrollcsoport átlaga közelített meg (82,5%, átlagos eltérés: 12,5%). Hibátlan teljesítményt a csoport 15%-a (3 gyermek) nyújtott; hatan-hatan 90%-ot, illetve 80%-ot, négyen 70%-ot értek el, míg egy gyermek az elhangzott értelmetlen hangsoroknak csupán a felét tudta megismételni. Az előző tesztekkel ellentétben itt a fiúk (84,6%, átlagos

eltérés: 10,6%) nyújtottak jobb teljesítményt (lányok átlaga: 76% átlagos eltérés: 16,7%).

A háromévesen beszélni kezdők átlagos teljesítménye 55% (átlagos eltérés: 24,4%) volt. 100%-os eredményt két gyermek ért el, egy-egy gyermek 80%-ot, 40%-ot, illetve 0%-ot nyújtott, hárman 70%-ot, és négyen-négyen 60%-ot, 50%-ot, illetve 30%-ot. A fiúk (átlag: 51%, átlagos eltérés: 21,2%) és a lányok (átlag: 66%, átlagos eltérés: 32,1%) értékei között 15 százalékpontnyi a különbség. Ebben a tesztben is a négyéves beszédindulású gyermekek teljesítettek a leggyengébben, csoportátlaguk 39,5% volt (átlagos eltérés: 18,5%). Náluk a legjobb eredmény 70%-nak adódott, amit csupán egyetlen gyermek ért el. Hárman-hárman a logatomok 60%-át, illetve 40%-át, hatan az 50%-át, öten a 20%-át, és ketten csupán a 10%-át tudták helyesen visszaadni, vagyis utóbbiak csak egyetlen hangsort. A fiúk (átlag: 38,6%, átlagos eltérés: 16,4%) és a lányok (átlag: 42%, átlagos eltérés: 25,9%) teljesítménye nem mutatott jelentős különbséget (5. ábra).



5. ábra

A szeriális észlelés (GMP10) eredményei a három csoportban (medián és szóródás)

A csoportok átlagteljesítménye szignifikáns különbséget mutatott ( $F(2, 446) = 21,994$ ;  $p = 0,001$ ). A páronkénti összehasonlítás igazolta, hogy mindegyik csoport különbözik a másik kettőtől (három, illetve négyévesen beszélni kezdők:  $t = 2,422$ ;  $p = 0,016$ ; háromévesen beszélni kezdők és kontrollcso-

port:  $t = 4,136$ ;  $p = 0,001$ ; négyévesen beszélni kezdők és kontrollcsoport:  $t = 6,558$ ;  $p = 0,001$ .

Elvégeztük az adatok páronkénti összevetését az elvárt értékekkel is, amely szerint minden csoport teljesítménye szignifikánsan elmarad az elvárt szinttől, még a kontrollcsoporté is (3 évesen beszélni kezdők:  $t = 7,632$ ;  $p = 0,001$ ; 4 évesen beszélni kezdők:  $t = 10,365$ ;  $p = 0,001$ ; kontrollcsoport:  $t = 2,965$ ;  $p = 0,003$ ).

### GMP12

A szövegértést vizsgáló tesztben hatéves korban 70-80% az elvárt teljesítmény a sztenderd értékeknek megfelelően. Ezt az eredményt a kontrollcsoportba tartozó gyermekek mindegyike elérte. A csoport átlagos teljesítménye 83%-nak (átlagos eltérés: 9,8%) adódott. A legtöbben – 9 gyermek – 80%-ot érték el, négyen-négyen 90%-ot, illetve 70%-ot, míg hárman hibátlanul válaszoltak minden megértést ellenőrző kérdésre. A lányok (átlag: 88%, átlagos eltérés: 8,4%) kevesebb, mint 7 százalékponttal teljesítettek jobban a fiúknál (átlag: 81,3%, átlagos eltérés: 9,9%).

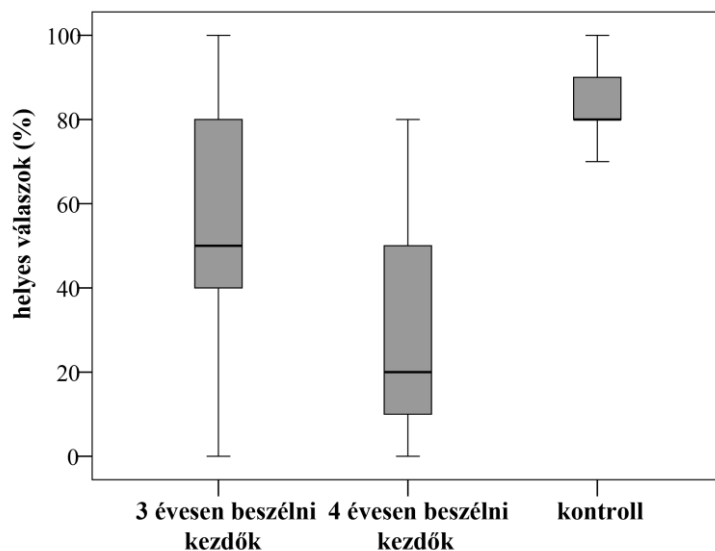
A megkésett beszédfejlődésű gyermekek teljesítménye ezen a területen is messze elmaradt mind a kontrollcsoportétól, mind az elvárt szinttől. Azok a gyermekek, akiknek a beszédindulása 3 éves korra tehető, 53%-os átlagteljesítményt (átlagos eltérés: 29,5%) értek el. Meglepetésre, ebben a csoportban 3 gyermek is 100%-os teljesítményt nyújtott, ketten-ketten 70%-osat, 80%-osat, illetve egyvalaki 90%-osat. 12 gyermek teljesítménye gyengébb volt; egy-egynél tudtunk 60%-ot, 10%-ot, illetve 0%-ot, két-két gyermeknél 50%-ot, illetve 30%-ot, és ötnél 40%-ot adatolni. A lányok átlagos teljesítménye 74% (átlagos eltérés: 37,8%), a fiúké mindössze 50% (átlagos eltérés: 24,78%) volt, ami a páronkénti összehasonlítás szerint szignifikáns különbség:  $t = 2,142$ ;  $p = 0,033$ .

A négyévesen beszélni kezdők a szövegértés terén is súlyos elmaradással küzdenek, átlagteljesítményük 29% (átlagos eltérés: 23,6%) volt. Az életkorban elvárt alsó értéket, a 80%-ot csupán egy gyermek érte el, ketten-ketten produkáltak 60%-ot, 40%-ot és 10%-ot, hárman 50%-ot, öten 20%-ot, négyen 0%-ot, és egy gyermek 30%-ot. A lányok (átlag: 30%, átlagos eltérés: 24,5%) és a fiúk (átlag: 28,7%, átlagos eltérés: 24,2%) teljesítménye nem mutatott eltérést. A statisztikai elemzés megerősítette, hogy a három csoport adatai szignifikánsan különböznek ( $F(2, 446) = 34,185$ ;  $p = 0,001$ ); az adatok páronkénti összevetése pedig mindhárom csoport között szignifikáns eltérést igazolt (a három- és a négyévesen beszélni kezdők csoportja között:  $t = 4,286$ ;  $p = 0,001$ ; a háromévesen beszélni kezdők és a kontrollcsoport között:  $t = 3,981$ ;  $p = 0,001$ ; illetve a négyévesen beszélni kezdők és a kontrollcsoport között:  $t = 8,267$ ;  $p = 0,001$ ). Az egyes csoportok eredményeit a 6. ábra szemlélteti. A statisztikai elemzés (egymintás  $t$ -próba) során a kapott adatokat a hatéves korban elvárt értékekkel is összevetettük. Eszerint mind a három- ( $t =$

2,111;  $p = 0,035$ ), mind a négyévesen beszélni kezdők ( $t = 6,947$ ;  $p = 0,001$ ), mind pedig a kontrollcsoport eredményei ( $t = 2,380$ ;  $p = 0,018$ ) matematikailag igazolhatóan különböztek az elvárt szinttől.

### GMP16

A rövid (mondatszerű) közlések szemantikai sajátosságainak és a grammatikai struktúráknak a feldolgozására fókuszál a mondatértés vizsgálata, amelyben hatéves korban 90-100%-os teljesítmény várható el. A kontrollcsoport átlagteljesítménye (84,5%, átlagos eltérés: 9,4%) mindössze 5,5 százalékponttal marad el az elvárttól. Három gyermek hibátlan, hat 90%-os, nyolc 80%-os, három pedig 70%-os teljesítményt nyújtott. A lányok (átlag: 86%, átlagos eltérés: 13,4%) és a fiúk (átlag: 84%, átlagos eltérés: 8,3%) eredménye nem különbözött jelentős mértékben egymástól.



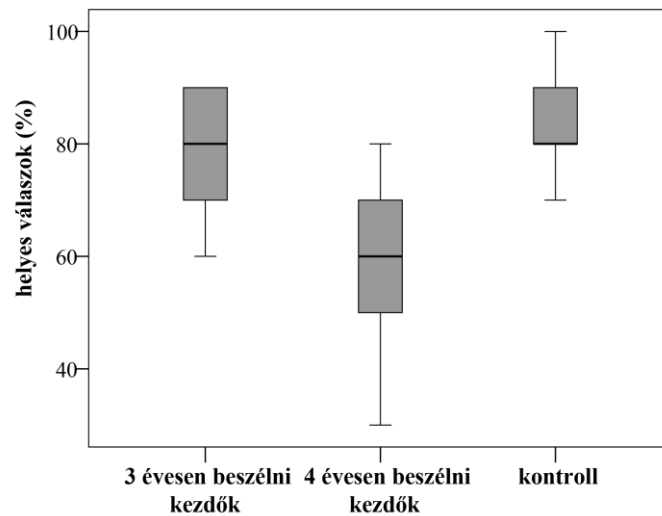
6. ábra

A szövegértés (GMP12) eredményei a három csoportban (medián és szóródás)

A háromévesen beszélni kezdő gyermekek csoportszinten 76%-ot (átlagos eltérés: 14,7%) értek el. Az elvárt szintet (90%) hat gyermek is teljesítette; további hat nyújtott 70%-os, öt 80%-os, kettő 60%-os, egy pedig 30%-os teljesítményt. Az eddig tapasztalt tendenciával ellentétben, ebben a csoportban

a fiúk valamivel eredményesebbek voltak (átlag: 77%, átlagos eltérés: 10,1%), mint a lányok (átlag: 70%, átlagos eltérés: 23,5%).

A négyévesen beszélni kezdők átlagosan a mondatok 59,5%-át (átlagos eltérés: 16,7%) tudták azonosítani; négyen-négyen a 80%-át, illetve 70%-át, hatan a 60%-át, ketten az 50%-át, hárman a 30%-át és egy gyermek a 40%-át. A lányok (átlag: 62%, átlagos eltérés: 10,9%) és a fiúk (átlag: 58,7%, átlagos eltérés: 18,5%) teljesítménye gyakorlatilag nem különbözött. A nemek között egyik csoportban sem volt kimutatható matematikailag igazolható különbség. A három csoport teljesítménye szignifikánsan különbözött ( $F(2, 446) = 7,400$ ;  $p = 0,001$ ). A páronkénti összehasonlítás szerint a 3 és a 4 évesen beszélni kezdők csoportja ( $t = 2,557$ ;  $p = 0,010$ ), illetőleg a 4 évesen beszélni kezdők és a kontrollcsoport között ( $t = 3,762$ ;  $p = 0,001$ ) áll fenn szignifikáns különbség (7. ábra). Az adatokat az elvárt szinttel összevetve megállapítható, hogy attól csak a három ( $t = 2,199$ ;  $p = 0,028$ ), illetve a négyévesen beszélni kezdők eredményei ( $t = 5,107$ ;  $p = 0,001$ ) különböznek matematikailag igazoltan; a kontrollcsoport teljesítménye megfelel az elvárásoknak.



7. ábra

A mondatértés (GMP16) eredményei a három csoportban (medián és szóródás)

### GMP17

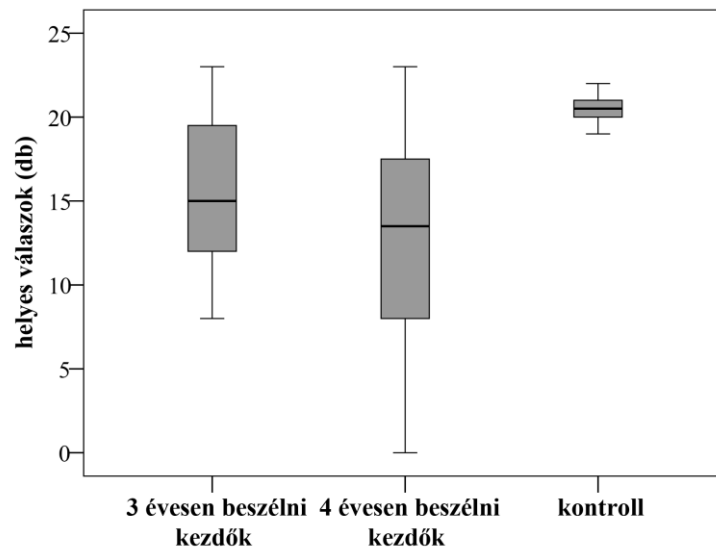
A beszédhang-differenciálás működését a teszt értelmetlen hangsorok ismételtetésével vizsgálja. 6 éves korban a mássalhangzók időtartam-

különbségének felismerésén kívül a gyermekeknek az összes hangorról el kell tudni dönteniük, hogy páronként azonosak vagy különbözők-e. Az elvárt helyes válaszok száma 19 db (az azonos logatomokat tartalmazó párokat is beleértve), ami 82,6%-os teljesítményt jelent. Az adatok szerint csak a kontrollcsoport átlaga érte el, sőt meg is haladta a kívánt szintet, ők átlagosan 20,4 db logatompárt (átlag: 88,7%, átlagos eltérés: 4,7%) tudtak megfelelően differenciálni. Megjegyezendő ugyanakkor, hogy csak ebben a csoportban fordult elő, hogy egyetlen gyermek sem ért el 100%-os eredményt. Hárman 22 (95,7%), illetve 19 (82,6%), heten 21 (91,3%), hatan 20 (87%) párról tudták helyesen eldönteni, hogy azok azonosak vagy különbözők voltak-e; egy gyermek eredménye (18 helyes találat, 78,3%) éppen elmaradt az elvárt értéktől. A fiúk (átlag: 88,4%, átlagos eltérés: 5,4%) és a lányok (átlag: 89,6%, átlagos eltérés: 2,2%) teljesítménye között gyakorlatilag nem volt különbség.

A háromévesen beszélni kezdők csoportszinten 15 logatompárt differenciáltak helyesen, ami 65,2%-os teljesítmény (átlagos eltérés: 25,1%); az adatok azonban széles skálán szóródnak. Két gyermek ért el hibátlan eredményt, kető-kettő 91,3%-ot, 52,2%-ot, illetve 34,8%-ot, hárman 65,2%-ot, egy-egy gyermek pedig 87%-ot, 82,6%-ot, 78,3%-ot, 73,9%-ot, 69,6%-ot, 60,9%-ot, 56,5%-ot, 39%-ot, illetve egyetlen hangsort sem volt képes helyesen megkülönböztetni. A fiúk és a lányok átlagteljesítménye egyaránt 65,2% (átlagos eltérés a fiúknál: 23,4%, átlagos eltérés a lányoknál: 32,8%) volt.

A négyévesen beszélni kezdők átlagosan a hangsortpárok 52,6%-át (átlagos eltérés: 30,6%) differenciálták megfelelően; az adatok szóródása ebben a csoportban is nagy. Három gyermek érte el vagy haladta meg az elvárt szintet (100%, 87%, 82,6%). További három gyermek minden hangsortpár esetében helytelenül döntött (0%), ketten-ketten 78,3%-ot, 73,9%-ot, 60,9%-ot és 56,5%-ot értek el. Egy-egy gyermeknél voltak adatolhatók a következő értékek: 65,2%, 52,2%, 47,8%, 39%, 30,4% és 8,7%. Az elvárt szinthez képest ismét csak a három- ( $t = 2,812$ ;  $p = 0,005$ ), illetve a négyévesen beszélni kezdők ( $t = 5,081$ ;  $p = 0,001$ ) mutattak elmaradást.

A lányok teljesítménye (átlag: 43,5%, átlagos eltérés: 26,2%) ebben a csoportban gyengébb volt, mint a fiúké (átlag: 55,7%, átlagos eltérés: 32,2%), de a különbség nem szignifikáns. A csoportok azonban szignifikánsan különböznek egymástól ( $F(2, 446) = 15,361$ ;  $p = 0,001$ ); a páronkénti összehasonlítás pedig a kontrollcsoport és a három- ( $t = 2,792$ ;  $p = 0,007$ ), illetve a négyévesen beszélni kezdők között ( $t = 4,602$ ;  $p = 0,001$ ) igazolt matematikailag szignifikáns különbséget. A 3 és a 4 évesek között nem volt szignifikáns eltérés kimutatható (8. ábra).



8. ábra

A beszédhang-differenciálás (GMP17) eredményei a három csoportban (medián és szóródás)

### GMP18

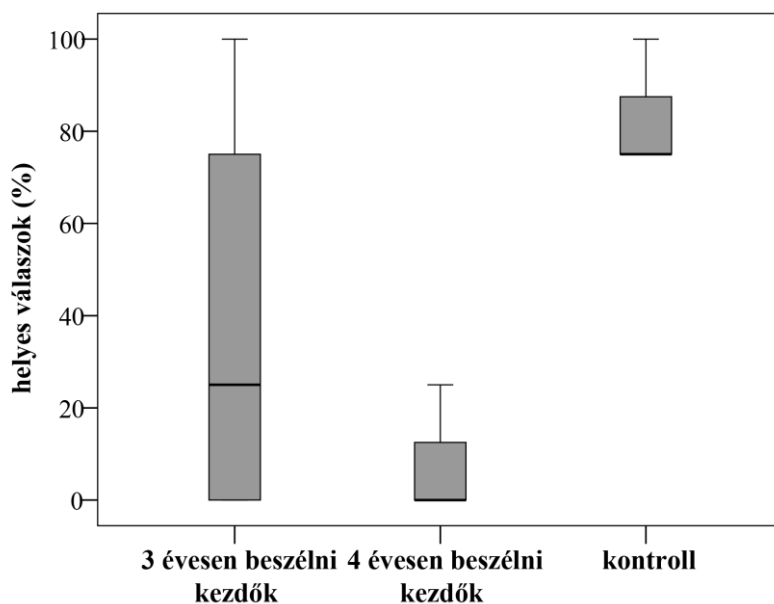
A transzformációs észlelés, vagyis az a folyamat, amely a beszédhang és a betű megfeleltetésének (későbbi) működését teszi lehetővé, meghatározó az írott anyanyelv megtanulásában. Hatéves korban 75%-os eredmény az elvárás ebben a tesztben. Ezt a szintet csak a kontrollcsoportba tartozó gyermekek érték el, átlagos teljesítményük 78,8% (átlagos eltérés: 14,7%) volt. Öt gyermek hibátlan eredményt ért el, tizenhárom 75%-osat, míg kettő 50%-osat. A lányok (átlag: 85%, átlagos eltérés: 13,7%) majdnem 9 százalékponttal teljesítettek jobban, mint a fiúk (átlag: 76,7%, átlagos eltérés: 14,8%).

A megkésett beszédfejlődésű gyermekek mindkét csoportja súlyos elmaradást mutatott. A háromévesen beszélni kezdők csoportátlaga 37,5% (átlagos eltérés: 37,6%) volt. Az elvárt szintet nyolc gyermek érte el, egy hibátlan, hét pedig 75%-os teljesítménnyel. 50%-ot egy, 25%-ot három gyermeknél, 0%-ot pedig nyolcnál adatoltunk. A lányok átlagos teljesítménye 40% (átlagos eltérés: 37,9%), a fiúké 37% (átlagos eltérés: 37,6%) volt.

A négyévesen beszélni kezdők átlagosan mindössze 11,3%-os eredményt (átlagos eltérés: 20,6%) értek el, a húsz gyermekből 15-en nem voltak képesek a teszt megoldására, egy gyermek 25%-ot és négy 50%-ot ért el. A lányok (átlag: 20%, átlagos eltérés: 27,4%) ugyan jobban teljesítettek, mint a



fiúk (átlag: 8,3%, átlagos eltérés: 18,1%), de az ő eredményük is messze elmarad az elvárttól (9. ábra).

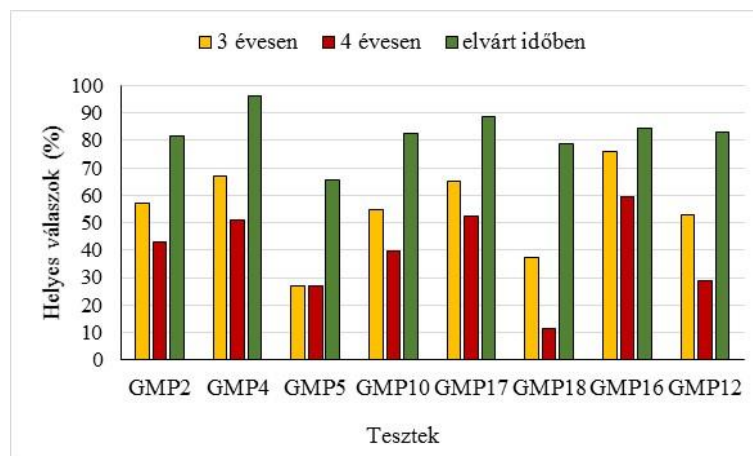


9. ábra

A transzformációs észlelés tesztjének (GMP18) eredményei a három csoportban (medián és szóródás)

A két nem között egyik csoportban sem volt kimutatható szignifikáns különbség. Az adatok eloszlása egyik csoportban sem véletlenszerű (a háromévesen beszélni kezdők esetében:  $\chi^2(4) = 11,000$ ;  $p = 0,027$ ; a négyévesen beszélni kezdőkénél:  $\chi^2(2) = 16,300$ ;  $p = 0,001$ ; a kontrollcsoportban:  $\chi^2(2) = 9,700$ ;  $p = 0,008$ ). A csoportok szignifikánsan különböztek egymástól ( $F(2, 446) = 54,388$ ;  $p = 0,001$ ). A páronkénti összehasonlítás a három- és a négyévesen beszélni kezdők ( $t = 4,169$ ;  $p = 0,001$ ), a háromévesen beszélni kezdők és a kontrollcsoport ( $t = 6,194$ ;  $p = 0,001$ ), valamint a négyévesen beszélni kezdők a kontrollcsoport között ( $t = 10,364$ ;  $p = 0,001$ ) igazolt szignifikáns különbséget (9. ábra). Az elvárt szinttől a statisztikai elemzések szerint csak a három- ( $t = 6,230$ ;  $p = 0,001$ ), illetve a négyévesen beszélni kezdők csoportja ( $t = 10,935$ ;  $p = 0,001$ ) különbözött szignifikánsan.

Az összes tesztben kapott átlageredményeket a csoportok szerint ábrázoltuk, így jól szemléltethetők a teljesítmények az adott beszédészlelési és beszédmegértési folyamat függvényében (10. ábra).



10. ábra

A vizsgált beszédfeldolgozási folyamatok eredményei a beszédindulás szerinti csoportokban (3 évesen, illetve 4 évesen, valamint az elvárt időben beszélni kezdők) és az egyes tesztekben

### Következtetések

Kutatásunk fő kérdése az volt, hogy vajon a beszédindulás időpontja milyen mértékben lesz meghatározó a vizsgált beszédfeldolgozási folyamatokra közvetlenül az iskolakezdés előtti életkorban (hatéveseknél). A megkésett beszédfejlődésű gyermekeket két csoportba osztottuk, az egyikbe azokat, akik háromévesen, a másikba azokat, akik négyévesen kezdtek beszélni (vagyis ezekben az életkorokban voltak a holofrázisok szakaszában).

Nemben és életkorban illesztett, tipikus beszédindulású hatévesek alkották a kontrollcsoportot. Minden teszt eredményeinek értékelésekor tekintetbe vettük a sztenderd értékeket is. Hat beszédészlelési folyamat működését, valamint a mondat- és a szövegértési teljesítményt elemeztük a GMP-diagnosztika (Gósy 1995/2006) alkalmazásával.

Az eredmények azt mutatták, hogy a beszédindulás ideje nagymértékben meghatározó az anyanyelv-elsajátítás folyamatára nézve a vizsgált receptív működések tekintetében. Mind a 3, mind a 4 évesen beszélni kezdők csoportja szignifikánsan elmaradt a kontrollcsoport teljesítményétől valamennyi tesztben. Megjegyezzük azonban, hogy adódott egy kivétel. A háromévesen beszélni kezdők mondatértési teljesítménye nem különbözött szignifikánsan a tipikus fejlődésűektől (a csoportátlagban 8,8 százalékpont volt közöttük az eltérés). (A csak 4 évesen szókezdeményeket mondóké azonban igen.) A 3 évesen beszélni kezdők és a tipikus fejlődésűek csoportjaiban a gyermekek egyéni teljesítményét is figyelembe véve jelentős eltéréseket tapasztaltunk. A

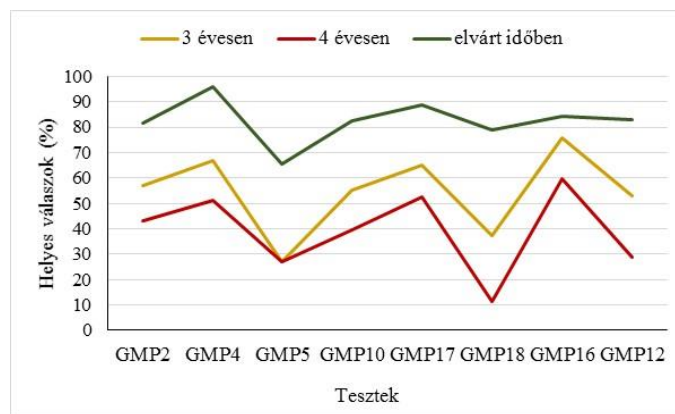
tipikus fejlődésűek 45%-a életkori vagy annál jobb (azaz hibátlan) teljesítményt nyújtott, további 40%-uk 80%-ot ért el a kívánt 90% helyett. A háromévesen beszélni kezdők 30%-a elérte az elvárt szintet, további 25%-uk pedig 8 mondatot értett meg jól (a minimálisan elvárt 9 helyett), a többi gyermek (45%) a csoportban jóval gyengébb eredményt mutatott. Az első hipotézisünk tehát – a mondatértés eredményétől kissé eltekintve – lényegében igazolódott.

Jelentős különbségeket adatoltunk attól függően is, hogy a kései beszédindulás három- vagy négyéves korban történt-e. A morfo-fonológiai észlelés és a beszédhang-differenciálás eredményeinek kivételével – ahol nem volt szignifikáns különbség a megkésett két csoportjának adatai között – minden tesztben matematikailag is igazolhatóan gyengébben teljesítettek a négyévesen beszélni kezdők. A morfo-fonológiai észlelés rendkívül gyengén működött valamennyi atipikus beszélő esetében, mindkét csoport átlagértéke 27% volt (az elvárt, életkori teljesítmény 90%). A beszédhangok differenciálása is szignifikánsan gyengébb volt mindkét csoportban, bár ebben a tesztben a 3 évesen beszélni kezdők kissé jobb átlageredményt értek el (62,5%-ot a másik késői beszédindulású csoport 52,6%-os átlagával szemben). A kontroll gyermekek adatainak átlaga 88,7%-nak adódott.

Megállapítható, hogy a 4 éves korban induló beszéd súlyosabb beszédfeldolgozási elmaradásokhoz vezet, mint a 3 évesen már megjelenő szókezdemények esetében. Lényegében tehát a második hipotézisünk is teljesült. A kései beszédindulás negatív hatással van a vizsgált feldolgozási folyamatokra; minél később kezd a gyermek az első szókezdeményeket ejteni, annál nagyobb elmaradások fogják jellemezni a beszédfeldolgozási szintjét hatéves korában.

Noha a GMP-diagnosztika gyakorlatában nem releváns az összes adat kimutatása egyetlen számban, itt most az összevetés érdekében megnéztük az egyéni teljesítményeket is ezen a módon. Azt tapasztaltuk, hogy a tipikus fejlődésűeknél a leggyengébben teljesítő gyermek 72%-ot, a legjobb 88%-ot ért el. A sztenderd értékek alapján elvárt érték 87%. A háromévesen beszélni kezdőknél a leggyengébb teljesítmény 18%, a legjobb 84% volt; a négyévesen beszélni kezdőknél pedig a leggyengébb eredmény 15%-nak, a legjobb 57%-nak adódott.

Harmadik hipotézisünkben azt feltételeztük, hogy a későn beszélni kezdők beszédfeldolgozási folyamatai különböző mértékű elmaradásokat fognak mutatni a vizsgált folyamattól függően. Ez a hipotézisünk is teljesült (vö. 11. ábra). Mind a háromévesen, mind a négyévesen beszélők hasonló mintázatot mutatnak az egyes folyamatokban nyújtott teljesítményük szerint. Az ábra az egyes tesztekben elért eredmények vonalas összekötésével ezt a mintázatot szemlélteti. Azoknál a gyermekeknél, akik négyéves korukban kezdtek beszélni, jellegzetesen nagyobbak az elmaradások, azaz gyengébb a csoport szintű teljesítmény.



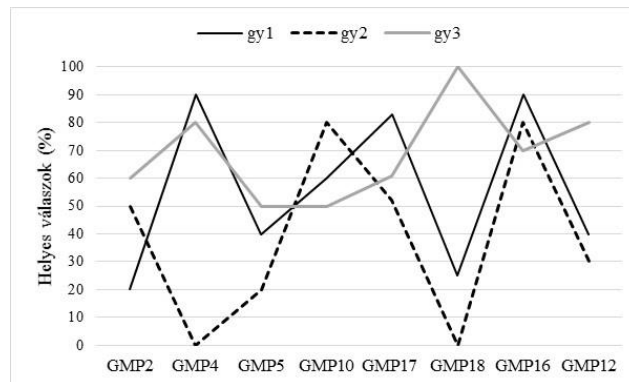
11. ábra

A vizsgált beszédfeldolgozási folyamatok eredményeinek mintázatai csoportonként és tesztenként (beszédindulás: 3 évesen, 4 évesen, illetve az elvárt időben)

A teljesítménymintázat egészét tekintve látható, hogy az hasonló a tipikus beszédfejlődésükhöz, ami azt jelenti, hogy a megkésett csoportok szintén egyértelműen elmaradást mutatnak és nem zavart (Gósy 1995/2006). Az elmaradások mértéke folyamatfüggő. Lényegesen kisebb például a mondatértésben, mint a transzformációs észlelésben. A csoportszintű teljesítménymintázat előjelzi azt, hogy a fejlesztés könnyebb lesz, mivel nem tapasztaltunk zavart, „csak” elmaradást. A gyermekek egyéni teljesítménye és percepciós mintázata azonban különbségeket mutat; és mindkét csoportban akadnak olyanok, akiknél nem egyszerű elmaradást, hanem percepciós zavart tapasztalunk, amint azt a 12. és a 13. ábra szemlélteti.

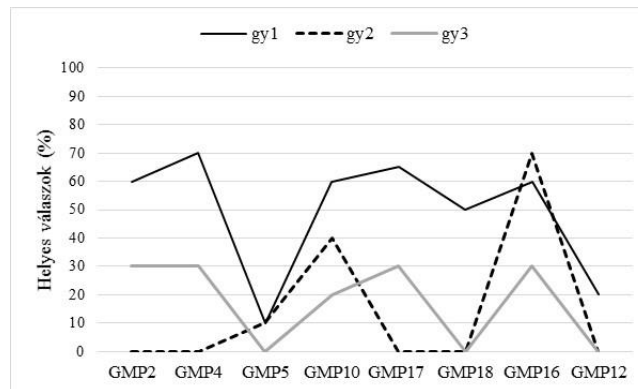
A késői beszédindulást mutató gyermekek beszédfeldolgozási folyamatainak elmaradása várható volt, ezt a nemzetközi és a hazai szakirodalom egyaránt megerősíti (pl. Duvelleroy-Hommet et al. 1995; Rescorla et al. 2000; Bishop–Snowling 2004; Gósy 2002a; Horváth 2007). A magyarázat önként adódik: a 3, illetve 2 év, ami alatt a gyermekek megpróbálják behozni a késésüket az anyanyelv-elsajátítás e kutatásban vizsgált területein, nem bizonyult elegendőnek. Mindez annak ellenére alakul így, hogy a megkésett beszédindulású gyermekek – kutatásunk résztvevői kivétel nélkül – logopédiai fejlesztést kaptak (illetve kapnak folyamatosan). A rendelkezésre álló idő mégsem volt elegendő ahhoz, hogy valamennyi, jelen kutatásban vizsgált folyamat elérje, avagy jól megközelítse az életkorban elvárt szintet. Ezek a gyermekek a mondatértésben produkálták a relatíve legjobb teljesítményt és a morfológiai, illetve a transzformációs észlelésben a leggyengébbet. Megjegyezzük, hogy ez utóbbiak működnek a legkevésbé jól a tipikus fejlődésük-

nél is. A négyévesen beszélni kezdők esetében a transzformációs észlelés és a szövegértés különösen jelentős elmaradást mutat.



12. ábra

A vizsgált beszédfeldolgozási folyamatok eredményeinek mintázatai tesztenként három gyermeknél (gy1, gy2, gy3), akik 3 évesen kezdtek beszélni



13. ábra

A vizsgált beszédfeldolgozási folyamatok eredményeinek mintázatai tesztenként három gyermeknél (gy1, gy2, gy3), akik 4 évesen kezdtek beszélni

A gyermekek további nyelvi fejlődésére nézve mindezek azért jelentenek nagy veszélyt, mert ezeknek a folyamatoknak a működése közvetlenül nem

tapasztható, nehezen megfigyelhető, vagyis a környezet nem feltétlenül veti észre az ebből fakadó nehézségeket. Hozzájárul ehhez az is, hogy a gyermekek egy része produkciós szinten nem mutat látványos eltérést a tipikus fejlődésűektől; és bár akadnak olyanok, akik beszédhibával küzdenek, de ez előfordul az egyébként tipikus anyanyelv-elsajátítási folyamatot mutató gyermekeknél is, és nem minden esetben jár együtt a beszédfeldolgozási működések elmaradásával (Gósy–Horváth 2015).

Ugyancsak a GMP-diagnosztika tesztjeivel vizsgáltak megkésett beszédfejlődésű gyermekeket mintegy 10 és 15 évvel ezelőtt (Gósy 2002c; Horváth 2007). Mindkét korábbi kutatásban 3 éves korukban beszélni kezdő gyermekek beszédpercepciósi folyamatait elemezték. Tekintettel a módszertan nagy hasonlóságára, (óvatosan bár) összehasonlítottuk a csoportátlagokat. A tendencia azt mutatja, hogy a jelen kutatásban részt vett megkésett beszédfejlődésű gyermekek tendenciaszerűen gyengébb csoportátlagokat mutattak, mint az egy évtizeddel, illetve másfél évtizeddel ezelőtt mértek. A GMP2-ben időrendben (Gósy 2002c; Horváth 2007; illetve a jelen kutatás) a háromévesen beszélni kezdő gyermekek helyes mondatismétlésének átlaga 68,3% volt, öt évvel később 59,2%, tíz évvel később pedig 57%. A GMP4-es teszt átlagai ugyanebben az időrendben: 77%, 73,6% és 67%. A GMP5-ben kissé módosulnak az egymáshoz viszonyított adatok: 26,7%, 36% és 27%. Itt a tíz évvel ezelőtti adatok javulást mutatnak a tizenöt évvel ezelőttiekhez képest, a jelen kutatás átlagértékei viszont gyakorlatilag megegyeznek a 2002-es tanulmány adataival. A szeriális észlelés (GMP10) csoportátlagai az időrendnek megfelelően: 46%, 54% és 55%. Itt tehát a leggyengébb eredményt a legkorábbi vizsgálatban adatolták, a tíz évvel ezelőtti és a jelen kutatás csoportátlagai nem mutatnak különbséget. Meg kell azonban jegyeznünk, hogy csupán a csoportátlagok összevetésére volt lehetőségünk, így nem állnak rendelkezésre a gyermekek egyéni teljesítményei. A fentiekben bemutatott különbségek, avagy egy részük az egyéni különbségekből is adódhat.

Az akusztikai-fonetikai és a fonetikai észlelés, illetve az akusztikai kulcsok felismerésének gyengülése a vizsgált populációban nyilvánvalóan sok tényezőre vezethető vissza. Ezen tényezők hatását e kutatásban nem vizsgáltuk, de a jövőben érdemes lenne az összefüggéseket kimutatni a késői beszédindulás ténye és például a gyermekek környezetében gyakran tapasztalható kevesebb verbális kommunikáció, a környezeti zajhatás, avagy a felnőttek beszédtempójának olykor relatíve gyors volta. Úgy tűnik, hogy ezeknek a hatásoknak az ellensúlyozása a célzott logopédiai fejlesztés során részlegesen sikerül.

A sorozatészlelés jól működik abban az értelemben, hogy itt javulás tapasztalható a korábbi csoportátlagokhoz képest. Feltételezhetően a terápia hatása ebben a folyamatban jobban megmutatkozik. A morfo-fonológiai folyamat (GMP5-ös teszt) gyengesége mindhárom vizsgálatban feltűnő, még ak-

kor is, ha a tíz évvel ezelőtti adatok közel 10 százalékponttal jobb eredményt mutattak.

A mondatértésben és a szövegértésben kapott adatokat a tíz évvel ezelőtti kutatás (Horváth 2007) eredményeivel tudtuk összevetni. Hangsúlyozzuk, hogy ez az összevetés több tekintetben is megkérdőjelezhető (nem ugyanazon gyermekek, nem ugyanazon tesztelési körülmények), mégis érdemes az adatokat szemügyre venni, mivel az azonos tesztfelvétel lehetőséget ad egyfajta óvatos összehasonlításra. A 10 évvel ezelőtti kutatásban a mondatértés csoportátlagai csaknem azonosak, 76,6% a korábbi és 76% a jelenlegi érték. A mondatok megértése – a diagnosztikában alkalmazott módszerrel mérve – úgy tűnik, biztos alapot nyújt a fejlesztéshez. Talán megkockáztatható az a kijelentés, hogy ebben a gyermekek teljesítménye változatlan az eltelt tíz év alatt. A szövegértésben pedig kifejezetten jobban teljesítettek a jelen kutatásban részt vett megkésett beszédfejlődésű gyermekek, a csoportátlaguk 53%, míg a tíz évvel ezelőtti adat 8,2 százalékponttal alacsonyabb, 44,8%. Mint-hogy az utóbbi évtizedben előtérbe került a gyermekek szövegértési problémája, úgy véljük, hogy a fejlesztés során erre fokozott hangsúly kerülhetett, aminek eredményeképpen a most vizsgált gyermekek szintje jobb lett, mint a korábban vizsgáltaké.

A tipikus fejlődésű gyermekek esetében is végeztünk néhány összevetést a tíz évvel ezelőtti szakirodalmi adatokkal (Bóna 2007; Markó 2007). A jelenlegi csoportátlagok kismértékű csökkenést mutattak a mondatértésben (4,3 százalékpontnyit), a korábbi csoportátlag 88,8%-nak (Markó 2007), a jelenlegi 84,5%-nak adódott. Ez lényegében nem jelent eltérést, különösen, ha az egyéni különbségeket is tekintetbe vesszük. A szeriális észlelésben lényegében nincs kimutatható változás a korábbi és a jelen adatok között; a tíz évvel ezelőtti átlagérték 81% (Bóna 2007), a jelenlegi pedig 82,5%. A többi összevethető adat azt jelzi, hogy a mostani értékek valamivel jobbak, mint a korábbiak voltak. Kismértékű pozitív változás tapasztalható a morfo-fonológiai észlelésben, 6,5 százalékpontnyi; 59% volt a korábbi átlagérték (vö. Bóna 2007) és 65,5% a mostani. Valamivel nagyobb mértékű az eltérés a szövegértésben; 8,4 százalékpont a különbség, ennyivel magasabb a csoportátlag a jelen kutatásban (83%) a korábbihoz viszonyítva (74,6%, vö. Markó 2007). A különbségek nem jelentősek, meglehetősen hasonló értékeket látunk a vizsgált csoportok tekintetében. Ezek az óvatos összevetések egyrészt azt mutatják, hogy a tipikus fejlődésű gyermekek teljesítménye többé-kevésbé állandónak mondható 10 év távlatában; másrészt úgy is fogalmazhatunk, hogy bár jelentős pozitív változást nem igazolnak az adatok, de nagymértékű gyengülés sem tapasztalható.

Minthogy az alkalmazott tesztek sztenderdizáltak, érdemes röviden kitérni a tipikus fejlődésű gyermekek teljesítményére a sztenderd értékekhez képest. Összességében az mondható, hogy a jelen kutatásban vizsgált csoport három folyamatban tükrözi az elvárt szintet, a beszédhang-differenciálásban

(GMP17), a transzformációs észlelésben (GMP18) és a szövegértésben (GMP12). Alig marad el a sztenderd értékektől a csoportátlag a fonetikai észlelésben (GMP4), illetve az akusztikai fonetikai percepcióban (GMP2), valamint a mondatértésben (GMP16). Jelentős az elmaradás ugyanakkor a morfo-fonológiai (GMP5) és a szeriális észlelés (GMP10) tekintetében, ami az érintett gyermekeknél rizikó tényezőnek minőséggül az írott anyanyelv megtanulására nézve.

Az anyanyelv elsajátításának korai szakaszai meghatározók az írott nyelv megtanulásának szempontjából (Desmarais et al. 2008; Preston et al. 2010; Perna–Loughan 2013). Úgy is fogalmazhatunk, hogy hosszú távú következményeket látunk az írás és az olvasás tanulására nézve azoknál a gyermekeknél, akiknél a beszédindulás nem az elvárt időben történt meg (Preston et al. 2010; Poll–Miller 2013). Adataink kétséget kizárólag alátámasztják ezeket a megállapításokat, ráadásul igazolják, hogy az egy év különbséggel beszélni kezdők között milyen nagymértékűek a különbségek. A négyévesen beszélni kezdő gyermekek olvasás- és írástanulása tehát már az iskoláskort megelőzően veszélyeztetett. A késői beszédindulás ténye erőteljes rizikótényező. A kutatásunkban részt vett gyermekek logopédiai fejlesztésben részesültek, így feltételezhető, hogy spontán beszédfejlődésük e nélkül a segítség nélkül még nagyobb elmaradásokat mutatna.

Eredményeink felvetik azt a megfontolást, hogy a beiskolázás időpontjának eldöntésekor szükséges lenne a megkésett beszédindulású gyermekek beszédfeldolgozási mechanizmusának ellenőrzése a későbbi nehézségek elkerülése érdekében, kivétel nélkül minden esetben. Adataink alapján úgy gondoljuk, a közvetlenül nem megítélhető beszédészlelési és beszédmegértési folyamatok működésének tesztelése a tipikus anyanyelv-elsajátítást mutató gyermekek esetében is ajánlott.

### Irodalom

- Aitchison, Jean 2003. *Words in the mind. An introduction to the mental lexicon*. Malden, Basil Blackwell.
- Berko Gleason, Jean – Bernstein Ratner, Nan (eds.) 1998. *Psycholinguistics*. Orlando, Harcourt Brace College Publishers.
- Billard, Catherine – Duvelleroy-Hommet, Caroline – de Becque, B – Gillet, P. 1996. Les dysphasies de développement [Developmental dysphasia]. *Archives de Pédiatrie* 3(6): 580–587.
- Bishop, Dorothy V. 2006. What causes specific language impairment in children? *Current Directions in Psychological Science* 15(5), 217–221.
- Bishop, Dorothy V. M. – Snowling, Margaret J. 2004. Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin* 130(6): 858–886.
- Bittera Tiborné – Juhász Ágnes 1991. *A megkésett beszédfejlődés terápiája*. Nemzeti Tankönyvkiadó, Budapest.



- Bóna Judit 2007. A fonológiai és a szeriális észlelés fejlődése 4–10 éves korban. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol, Budapest, 262–270.
- Boyle, Coleen A. – Boulet, Sheree – Schieve, Laura A. – Cohen, Robin A. – Blumberg, Stephen J. – Yeargin-Allsopp, Marshalyn – Visser, Susanna – Kogan, Michael D. 2011. Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics* 127. 1034–1042.
- Brooks, Patricia J. – Kempe, Vera (eds.) 2014. *Encyclopedia of language development*. SAGE Publications, Inc., Thousand Oaks.
- Chen, Yuchun – Tsao, Feng-Ming – Liu, Huei-Mei 2016. Developmental changes in brain response to speech perception in late-talking children: A longitudinal MMR study. *Developmental Cognitive Neuroscience* 19:190–199.
- Chiat, Shula – Roy, Penny 2013. Early predictors of language and social communication impairments at ages 9–11 years: A follow up study of early referred children. *Journal of Speech, Language, and Hearing Research* 56/6. 1824–1836.
- Chomsky, Noam 1995. *Mondattani szerkezetek – Nyelv és elme*. Osiris–Századvég, Budapest.
- ContiRamsden, Gina – Botting, Nicola – Faragher, Brian 2001. Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry* 42/6. 741–748.
- Crystal, David – Michael Garman – Paul Fletcher 1989. The grammatical analysis of language disability. Whurr Publishers Ltd., London.
- Csépe Valéria 2007. Azonos vagy különböző? Beszédészlelési és olvasási zavarok: diszlexia és SLI. In: Gósy M (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 164–183.
- Dale, Philip S. – Price, Thomas S. – Bishop, Dorothy V. M. – Plomin, Robert 2003. Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research* 46. 544–560.
- Desmarais, Chantal – Sylvestre, Audette – Meyer, François – Bairati, Isabelle – Rouleau, Nancie 2008. Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language Communication Disorders* 43/4. 361–389.
- Duvelleroy-Hommet, Caroline – Gillette, P. – Billard, Catherine – Loisel, M. – Barthez, M. A. – Santini, J. J. – Autret, A. 1995. Study of unilateral hemisphere performance in children with developmental dysphasia. *Neuropsychologia* 33. 823–834.
- Ellis Weismer, Susan 2007. Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In Paul, Rhea (ed.): *The influence of developmental perspectives on research and practice in communication disorders: A Festschrift for Robin S. Chapman*. Mahwah, NJ: Erlbaum. 83–102.
- Gerebenné Várbíró Katalin (szerk.) 1995. *Fejlődési diszfázia*. Bárczi Gusztáv Gyógypedagógiai Főiskola, Budapest.
- Gósy Mária 1995/2006. *GMP-diagnosztika*. Nikol, Budapest.

- Gósy, Mária 1998. Speech perception and comprehension of dysphasic children. In Ziegler, Wolfram – Deger, Karin (eds.): *Clinical linguistics and phonetics*. Whurr Publishers. London, 29–36.
- Gósy Mária 2002a. A diszfázia beszédpercepciósi tipológiája. *Pszichológia* 3. 277–289.
- Gósy Mária 2002b. A beszédészlelés és a beszédmegértés folyamatának zavarai. In Martonné Tamás Márta (szerk.): *Fejlesztő pedagógia*. ELTE Eötvös Kiadó, Budapest. 117–139.
- Gósy, Mária 2002c. Speech perception deficits and the underlying nature of developmental dysphasia. *Acta Linguistica Hungarica* 49/3-4. 363–383.
- Gósy Mária 2005. Pszicholingvisztika. Osiris Kiadó, Budapest.
- Gósy Mária – Horváth Viktória 2006. A percepciósi folyamatok összefüggései hatéveseknél. *Alkalmazott Nyelvtudomány* VI. 25–42.
- Gósy, Mária – Horváth, Viktória 2015. Speech processing in children with functional articulation disorders. *Clinical Linguistics and Phonetics* 29/3. 185–200.
- Horváth Viktória 2007. Megkésett beszédfejlődésű óvodások beszédfeldolgozási folyamatairól. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol, Budapest. 149–162.
- Klem, Marianne – Melby-Lervåg, Monica – Hagtvet, Bente – Halaas Lyster, Solveig-Alma – Gustafsson, Jan-Eric – Hulme, Charles 2015. Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science* 18/1. 146–154.
- Kovac, Ilija – Berdij Garabedian – Christele Du Souich – Roberta M. Palmour 2001. Attention deficit/hyperactivity in SLI children increases risk of speech/language disorders in first-degree relatives. A preliminary report. *Journal of Communication Disorders* 34. 339–354.
- Kraus, Nina – Therese J. McGee – Dawn Burton Koch 1999. Biology of speech sound perception and linguistic experience. In Maassen, Ben – Groenen, Paul (eds.): *Pathologies of speech and language*. Whurr Publishers Ltd., London, 117–137.
- Ludlow, Christy L. 1980. Children's language disorders: Recent research advances. *Annales Neurologica* 7. 497–507.
- Lyytinen, Paula – Eklund, Kenneth – Lyytinen, Heikki 2005. Language development and literacy skills in late-talking toddlers with and without familial risk for dyslexia. *Annals of Dyslexia* 55/2. 166–192.
- Marchman, Virginia A. – Fernald, Anne 2013. Variability in real-time spoken language processing in typically developing and late-talking toddlers. In Rescorla, Leslie A. – Dale, Philip S. (eds.): *Late talkers: Language development, interventions, and outcomes*. Baltimore, MD: Brookes, 145–166.
- Markó Alexandra 2007. A mondat- és szövegértés jellemzői és összefüggése 6–9 éves korban. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol, Budapest. 285–300.
- Martin, John A. M. 1981. *Voice, speech, and language in the child: Development and disorder*. Springer, Wien, New York.
- Miniscalco, Carmela – Nygen, Gudrun – Hagberg, Bibbi – Kadesjö, Björn – Gillberg, Christopher 2006. Neuropsychiatric and neuro-developmental outcome of children at age 6 and 7 years who screened positive for language problems at 30 months. *Developmental Medicine and Child Neurology* 48. 361–366.

- Palmour, Roberta M. 1997. Genetic studies of specific language impairment. *Journal of Neurolinguistics* 10. 215–230.
- Paul, Rhea – Murray, Candace – Clancy, Kathleen – Andrews, David 1997. Reading and metaphonological outcomes in late talkers. *Journal of Speech, Language, and Hearing Research* 40. 1037–1047.
- Paul, Rhea – Shiffer, M. E. 1991. Communicative initiations in normal and latetalking toddlers. *Applied Psycholinguistics* 12. 419–433.
- Perna, Robert – Loughan, Ashlee R. 2012. Early developmental delays: Neuropsychological sequelae and subsequent diagnoses. *Applied Neuropsychology: Child* 1. 57–62.
- Perna, Robert – Loughan, Ashlee R. 2013. Early developmental delays: A cross validation study. *Journal of Psychological Abnormalities in Children* 2013. 1/2. 121–126.
- Poll, Gerard H. – Miller, Carol A. 2013. Late talking, typical talking, and weak language skills at middle childhood. *Learning and individual differences* 26. 177–184.
- Preston, Jonathan L. – Frost, Stephen J. – Mencl, William Einar – Fulbright, Robert K. – Landi, Nicole – Grigorenko, Elena – Jacobsen, Leslie – Pugh, Kenneth R. 2010. Early and late talkers: schoolage language, literacy and neurolinguistic differences. *Brain* 133. 2185–2195.
- Psarommatis, Joannis M. – Eleni Goritsa – Dimitrios Douniadakis – Michael Tsakanikos – Alexandra D. Kontogianni – Nikolaos Apostolopoulos 2001. Hearing loss in speech-language delayed children. *International Journal of Pediatric Otorhinolaryngology* 58. 205–210.
- Rescorla, Leslie – Dahlsgaard, Katherine – Roberts, Julie 2000. Late talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. *Journal of Child Language* 27. 643–664.
- Rescorla, Leslie 2002. Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research* 45. 360–371.
- Rescorla, Leslie A. 2005. Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research* 48. 459–472.
- Rescorla, Leslie – Turner, Hannah H. 2015. Morphology and syntax in late talkers at age 5. *Journal of Speech, Language, and Hearing Research* 58/2. 434–444.
- Rice, Mabel L. – Redmond, Sean M. – Hoffman, Lesa 2006. Mean length of utterance in children with SLI and younger control children shows concurrent validity, stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research* 49. 793–808.
- Rice, Mabel L. 2007. Children with specific language impairment: Bridging the genetic and developmental perspectives. In Hoff, Erika – Shatz, Marilyn (eds.): *Blackwell handbook of language development*. Malden, MA: Blackwell. 411–431.
- Rice, Mabel L. – Taylor, Catherine L. – Zubrick, Stephen R. 2008. Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research* 51. 394–407.
- Roos, Elizabeth Michelle – Ellis Weismer, Susan 2008. Language outcomes of late talking toddlers at preschool and beyond. *Perspectives of Language Learning and Education* 15/3. 119–126.

- Rosta Katalin – Schuchné Rumpli Henrietta 2007. Beszédfogyatékos gyermekek beszédpercepció vizsgálat. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol, Budapest. 137–148.
- Rudolph, Johanna M. – Leonard, Laurence B. 2016. Early language milestones and specific language impairment. *Journal of Early intervention* 38/1. 41–58.
- Scarborough, Hollis S. 1990. Index of productive syntax. *Applied Psycholinguistics* 11. 1–22.
- Scarborough, Hollis S. – Dobrich, Wanda 1990. Development of children with early language delay. *Journal of Speech and Hearing Research* 33. 70–83.
- Seeff-Gabriel, Belinda – Chiat, Shula – Dodd, Barbara 2010. Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language and Communication Disorders* 45/6. 691–702.
- Thal, Donna J. – Tobias, Stacy – Morrison, Deborah 1991. Language and gesture in late talkers: A one year followup. *Journal of Speech, Language, and Hearing Research* 34/3. 604–612.
- Tallal, Paula – Sainburg, Robert – Jernigan, Terry 1991. The neuropathology of developmental dysphasia: Behavioral, morphological, and physiological evidence for a pervasive temporal processing disorder. *Reading and Writing* 3/3-4. 363–377.
- Wyke, Maria A. (ed.) 1978. *Developmental dysphasia*. Academic Press Inc., London–New York.
- Zubrick Stephen R. – Taylor Catherine L. – Rice Mabel L. – Slegers, David W. 2007. Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research* 50. 1562–1592.
- Zsoldos Márta – Csányi Yvonne 2010. Diszfáziás hallássérült tanulók (gyógy)pedagógiai tipológiája. *Gyógypedagógiai Szemle* 1. [http://www.prae.hu/prae/gyosze.php?menu\\_id=102&jid=31&jaid=453](http://www.prae.hu/prae/gyosze.php?menu_id=102&jid=31&jaid=453) (Letöltés ideje: 2017. jún. 24.)

### Receptive skills of children with late language onset

Late language emergence is diagnosed when language development trajectories are below age expectations with no diagnosed disabilities or developmental delays in other cognitive or motor domains. Some children with late talking onset are reported to show normal range receptive language abilities while others have various impairments in receptive language skills. The question arises whether the onset of language emergence influences children's receptive skills assessed at the age of 6. The aim of this study was to investigate the receptive skills of two groups of Hungarianspeaking preschool children whose language onset was either at the age of 3 or at the age of 4. Twenty children with a history of language onset at 3 years and another twenty children with a history of language onset at 4 years were examined. Their receptive skills were compared to a group of twenty typically developing children matched on age, articulation, and

socioeconomic status. For assessment, six subtests (three types of sentence repetition, nonword repetition, speech sound discrimination, text comprehension) of the Hungarian GMP testpackage were used in individual sessions.

As expected, group means for children with language onset at 4 were significantly below those for children with language onset at 3. Both groups underperformed significantly as compared to the typically developing children's performance. The results confirm that children with late language onset have serious impairments in morphophonological processing and text comprehension, foreboding difficulties in learning to read and write.

## DISFLUENCIES AND DISFLUENCY CLUSTERS IN CLUTTERED, STUTTERED AND TYPICAL SPEECH

**Judit Bóna**

ELTE, Department of Phonetics

### **Introduction**

Cluttering and stuttering are fluency disorders which can appear separately or together, too. Cluttering “is a fluency disorder wherein segments of conversation in the speaker’s native language typically are perceived as too fast overall, too irregular, or both. The segments of rapid and/or irregular speech rate must further be accompanied by one or more of the following: (a) excessive ‘normal’ disfluencies; (b) excessive collapsing or deletion of syllables; and/or (c) abnormal pauses, syllable stress, or speech rhythm” (St. Louis–Schulte 2011: 241–242). According to one of the most widely used definition from a behavioral standpoint (Wingate 1964: 488), stuttering is a “(a) Disruption in the fluency of verbal expression, which is (b) characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements, namely: sounds, syllables, and words of one syllable. These disruptions (c) usually occur frequently or are marked in character and (d) are not readily controllable.” The definition of the World Health Organization listing (2016) is similar: “Speech that is characterized by frequent repetition or prolongation of sounds or syllables or words, or by frequent hesitations or pauses that disrupt the rhythmic flow of speech. It should be classified as a disorder only if its severity is such as to markedly disturb the fluency of speech”. (ICD 10, 98.5)

Although, according to the Lowest Common Denominator definition (St. Louis–Schulte 2011: 241–242), cluttering can be characterized by excessive disfluencies, there are only a few papers about the types, characteristics, and the frequency of disfluencies in cluttered speech. These papers (St. Louis et al. 1985; Myers–St. Louis 1996; Van Zaalen et al. 2009a; Oliveira et al. 2010; Myers et al. 2012) compare disfluencies in cluttering with those in stuttering or in typical speech. Namely, the examination of the occurrences of disfluencies might be applicable in the differential diagnosis between cluttering and stuttering (Myers et al. 2012).

Disfluencies occur in the speech of typical speakers, too (Levelt 1989; Lickley 2015). Typical speech and fluency disorders might be distinguished from each other by the frequency and types of disfluencies. In addition, com-

paring typical speech and fluency disorders, there might be differences in speech rate as well. Speech rate in cluttering might be faster (Bakker et al. 2011), while in stuttering slower (Andrade et al. 2003) than in typical speech. However, there are differences in the disfluencies of cluttering and stuttering, too. According to the speech therapy literature, there are two main types of disfluencies: typical and stuttering-like disfluencies. Both types could occur in the speech of every speaker: there are stuttering-like disfluencies in typical speech and typical disfluencies in stuttered speech (Tetnowski–Scott 2010). The difference between the two types of speech is in the incidence rate. In English, the most frequent disfluencies are the same in cluttering and typical speech (Myers et al. 2012). The rate of stuttering-like disfluencies in cluttered and typical speech is only 2-3% (Myers et al. 2012). In stuttering, the occurrence of disfluencies is more frequent and most of them are stuttering-like (Tetnowski and Scott (2010). Table 1 shows the types of disfluencies with regards to indication whether they are stuttering-like or typical.

9. Table 1: Types of disfluencies based on Roberts et al. (2009) and Tetnowski and Scott (2010)

Disfluency type	Definition	Example	Typical or stuttering-like disfluency
<b>Interjection</b>	Any sound, syllable or extraneous word which does not contribute in meaning to the sentence.	The most frequent filler words in English: <i>well, like, you know</i> . The most frequent filler sounds: <i>um, uh</i> .	Typical disfluency
<b>Part-word repetition</b>	A sound or syllable said more than once with no additional meaning.	<i>I've been playing p-p-p-piano for a long time.</i>	Stuttering-like
<b>Word repetition</b>	A word said more than once with no additional meaning.	<i>There is a book on the the table.</i>	Typical disfluency (The repetition of one-syllable words are considered as stuttering-like in further studies. However, it is mostly typical in adult speakers' speech.)
<b>Phrase repetition</b>	More than one word said more than once with no	<i>I think I think you are right.</i>	Typical disfluency

	additional meaning.		
<b>Revision</b>	Instances when the speaker corrects an error.	<i>I see <b>the hor- the dog.</b></i>	Typical disfluency
<b>Incomplete phrase</b>	The speaker begins but does not complete an utterance.	<i><b>He is-</b> oh I forgot where he works.</i>	Typical disfluency
<b>Broken word</b>	A pause within the word.	<i>I <b>for</b>[pause]got his name.</i>	Stuttering-like (In agglutinative languages like Hungarian, broken word might be typical disfluency. The pause within the word might show e.g. difficulty in linguistic formulation, if the pause occurs between the root of the word and the suffix, Gósy 2012a.)
<b>Prolonged sounds</b>	Any sounds considered longer than normal.	<i>Ssssssssssummer was too hot.</i>	Stuttering-like
<b>Tense pause (block, stoppage, fixation)</b>	There are inaudible or hardly audible manifestations of muscular tension or tense articulatory contact between words, part-words, non-words.	... [pause with tension] I saw.	Stuttering-like

As regards to frequency, Oliveira et al. (2010) when analysing the speech of Brazilian Portuguese speakers, found that persons who clutter (PWC) produce twice as many disfluencies than typical speakers. This was typical of the case of both stuttering-like and other disfluencies. They found more than twice as many interjections, revisions and unfinished words, and seven times more word repetitions in the speech of PWC than in that of the typical speakers. However, Bakker et al. (2011), Myers et al. (2012), and Bóna (2016) found that there are not any significant differences between PWC and typical speakers when examining the frequency of disfluencies. Myers et al. (2012) analysed the occurrence and types of single disfluencies and disfluency clusters. Comparing the speech of PWC and typical speakers, they did not find any differences in either the frequency, or the most frequent types of disflu-



encies. There was significant difference between the two groups only in the occurrence of revisions in clusters and word repetitions in clusters. It may occur that in certain speech tasks (for example in rhetorical speech in front of audience, Bóna 2012), typical speakers produce more disfluencies than PWC.

Van Zaalen et al. (2009a) found that the speech task largely determines the proportion of typical and stuttering-like disfluencies. They assumed that PWC do not have enough time for speech planning which causes the occurrence of disfluencies in cluttering (van Zaalen et al. 2009b). This is also proven by the fact that when they slow down their own speech, less disfluencies occur (Bóna 2012). According to the Cluttering Spectrum Hypothesis (Ward 2006) there is a big overlap between the characteristic features of typical and PWC's speech. Typical speakers might produce symptoms of cluttering, too. Since cluttering is a multidimensional disorder, there are other factors (fast or irregular speech rate, poor speech intelligibility, inappropriate prosody etc.) in addition to the frequency of disfluencies which might influence its perception. The occurrence of excessive disfluencies is a possible but not obligatory characteristic feature of cluttering (St. Louis–Schulte 2011: 241–242; Myers et al. 2012).

There are various explanations for the occurrence of disfluencies in stuttering. On the one hand, stuttering-like disfluencies are disfluencies which appear on the motoric level (Ward 2006). This means that the speaker knows what they want to say but disruptions are in the motoric sequencing of sounds and syllables. According to the Covert Repair Hypothesis (Postma–Kolk 1993) people who stutter (PWS) have weak phonological encoding. This is why there are many disfluencies in their inner speech. The part-word repetitions in overt speech are consequences of covert self-monitoring. According to the Vicious Circle Hypothesis (Vasic–Wijnen 2001), PWS do not have weak phonological encoding. The problem is that PWS consider their own speech atypical. This is why the monitor becomes hypervigilant and this leads to excessive disfluencies.

Disfluencies often form clusters (Hubbard–Yairi 1988; LaSalle–Conture 1995). This means that two disfluencies appear in the same word or adjacently. Examples for two-elements disfluency clusters are the following (from LaSalle–Conture 1995): *I-I-I-you were going, I was I was um going*.

The frequency of disfluency clusters carries diagnostic value: it might show the severity of stuttering (Hubbard–Yairi 1988; LaSalle–Conture 1995) and cluttering (Myers et al. 2012). Analysing the speech of two PWC, Myers et al. (2008) found that PWC with more severe cluttering produced four times as many clusters as PWC with less severe cluttering. The types of disfluencies were not influenced by the severity of cluttering. The most frequent cluster constituents were interjections, revisions and unfinished words. Myers et al. (2012) compared the speech of 18 PWC and 20 typical speakers. Accord-

ing to their results, PWC produced much more disfluency clusters than typical speakers. However, the types of clusters were the same in both groups.

Occurrences of disfluency clusters were analysed in stuttering, too. According to the results of LaSalle and Conture (1995), children with stuttering (CWS) produced significantly more stuttering-stuttering clusters than other types. Typical children never produced stuttering-stuttering clusters. Stuttering-repair clusters (which contain a stuttering-like disfluency and a repair) occurred significantly more times in the speech of the former group. Repair-repair clusters (which contain two repairs) occurred more frequently in the speech of the latter group. Comparing the speech of adults who stutter (AWS) and the speech of CWS, the following difference was found: the ratio of disfluency clusters in the speech of CWS was more than half of all disfluency types (Hubbard–Yairi 1988; LaSalle–Conture 1995), while in the speech of AWS this ratio was no more than one third of the total number of disfluencies (Robb et al. 2009).

Disfluency clusters in which there are more than two disfluencies together, one right after the other, might also occur. These disfluency clusters are often called complex disfluencies. They are a series of disfluencies in succession (Shriberg 1994; Heeman et al. 2006; Robb et al. 2009; Gósy 2012c). Disfluency clusters containing more than two elements show bigger speech planning and production problems in both typical and atypical speakers.

There are many studies analysing the occurrence of disfluency clusters in stuttering, but there are only a few papers about disfluency clusters and disfluencies in general in cluttering. The aim of this paper is to examine the differences among cluttering, stuttering, and typical speech with regards to the frequency and types of disfluencies – both singletons and disfluency clusters. The frequency of complex disfluencies which contain more than two elements will be analysed, too.

According to the first hypotheses (1) the greatest difference will be between typical speakers and PWS in the frequency of the different types of disfluencies. (2) There will not be difference between typical speakers and PWC in the frequency of disfluencies. (3) PWS and PWC produce significantly more complex disfluencies than typical speakers.

## **Methodology**

### **Participants**

21 speakers participated in the study. There were 7 PWC, 7 PWS and 7 typical speakers. PWC were recruited by speech therapists for the study. Speech samples of PWS group were taken from the speech corpora of Bóna (2008), Gósy–Bóna (2011) and Kántor (2015). Control speakers were university students who participated in the recordings of BEA Hungarian Speech Database (Gósy 2012b). All participants volunteered for the tasks. Two experts in fluency disorders, a speech-language pathologist and a linguist spe-

cialized in fluency disorders, determined the diagnostic decisions independently of each other. They classified the speakers based on the recordings in three groups. Their rate of agreement was 100%.

In the group of PWC, speakers were aged between 20 and 32. There were 6 males and 1 female. All of them were native Hungarian speakers with normal hearing, and they had at least 14 years of education. They were pure clutterers, and they did not have any comorbid speech, language, cognitive, or psychiatric disorders. They did not have a history of stuttering in the past. All of them were aware of their speech problems and they considered themselves to be PWC. Based on the judgements of the two experts, subjects were classified as PWC if they had perceptually rapid and/or irregular speech rate and if their speech was characterized at least by one of the following (Bakker et al., 2011): (1) perceptually excessive disfluencies (the majority of them were non stuttering-like; the judgement was based on the subjective impression of the two experts), and/or (2) specific articulation characteristics which manifested in coarticulated speech or omissions of sounds and syllables, which excluded dyslalia or any other articulation disorder (Van Zaalen-op 't Hof et al., 2009a; Bakker et. al., 2011). This latter means that they produced some words like the following: *apveten* 'alapvetően, basically', *örkös* 'örökös, continual', *szeretnék* 'szeretnék, I would like'. Not everyone's speech was characterized by both properties.

In the group of PWS, speakers were aged between 20 and 56 (six speakers were between 20 and 32, one speaker was 56-year-old). There were 5 males and 2 females. All of them were native Hungarian speakers with normal hearing, and they had at least 14 years of education. They had no hearing, neurological, mental, speech or language deficits. There were 4 PWS with severe stuttering, and 3 PWS with mild stuttering. The severity of stuttering was determined on the basis of a questionnaire for diagnosis (Feketéné Gacsó-Mácsainé 2007), and on subjective judgements on the frequency of disfluencies. All PWS participated in speech therapy.

In the control group, speakers were aged between 20 and 32. There were 6 males and 1 female. All of them were native Hungarian speakers with normal hearing, and they had at least 14 years of education.

### Material

Spontaneous narratives were recorded from each speaker about the same topic: they were asked to speak about their own education, work, hobbies, and family. The interviewers let the participants talk freely, and they asked a question only when the speaker didn't know how to continue. The recordings were made digitally in a soundproof chamber. Analysis of 200 syllables of speech from each speaker was carried out according to the recommendation of the Systematic Disfluency Analysis (SDA) (Campbell-Hill 1994). In the

analysed Hungarian speech samples, 200 syllables were equal to 108-120 words.

### Method

The occurrence of disfluencies were coded in the speech samples. The types of disfluencies were the following: interjections, whole-word repetitions, part-word repetitions, broken words, prolongations, revisions. Interjections were counted in total and in two types separately, too. (1) The first type contained fillers which were pauses filled with sounds (not words) (Fletcher 2010). In Hungarian, the most frequent filled pauses are *ö*, *m* *öm*, *öh* (Horváth 2010). (2) The second type contained filler words which meant the interjection of real words.

Disfluencies were considered singletons if there was only one disfluency in a word or between two fluent words. Disfluencies were considered disfluency clusters if more than one disfluencies occurred on the same word or adjacently. For example, disfluency clusters occurred in repetitions with a prolongation on one of the repeated elements, or with a filler between the elements. Prolongation on interjections (filled pauses) were not counted (as in Roberts et al. 2009). When there were two interjections between two fluent words (for example: *hát öö* 'well uhm') the case was considered as a cluster.

Calculating the frequency in clusters containing more than two elements (when comparing disfluencies as singletons to disfluencies in clusters), every occurrence was counted in the given type of disfluency cluster. After this comparison, the clusters which contained more than two disfluencies, i.e. complex disfluencies, were separately examined. Frequency values show how many disfluencies occurred in the analysed 200 syllables. All calculations and ratings were carried out twice by the author, two weeks apart. The results of the two analyses were similar in 100% of the cases.

Statistical analysis (Mann-Whitney test, Wilcoxon test, Pearson-correlation) was carried out by SPSS in a 95% confidence level.

### Results

Table 2 shows the frequency of disfluencies by type and in total, and if they occurred as a singleton or in clusters in the speech of typical speakers, PWC and PWS. There was significant difference between typical speakers and PWS, and PWC and PWS in the frequency of all disfluencies (Table 3). PWS produced disfluencies more frequently than any of the other two groups. There was no significant difference between typical speakers and PWC in the frequency of all disfluencies.

There was no significant difference between the groups in the frequency of disfluency clusters despite the big differences in the means due to the large individual deviations in the background. For example, one speaker of the

PWS group produced only 6 disfluencies in clusters, while in the speech of another PWS with severe stuttering, 85 disfluencies in clusters occurred.

Table 2: Frequency of occurrence of disfluency types in 200 syllables (s = occurring as singletons, cl = occurring in clusters)

	Typical		PWC		PWS	
	Mean	SD	Mean	SD	Mean	SD
<b>Interjections</b>						
s	6.14	2.41	3.29	2.06	4.57	2.94
cl	5.14	3.93	5.86	5.46	18.29	24.24
<b>Total</b>	11.28	5.53	9.15	6.69	22.86	24.87
<b>Fillers</b>						
s	5.71	2.29	1.86	1.86	4.57	2.94
cl	4.00	3.16	3.71	3.20	17.00	23.52
<b>Total</b>	9.71	4.92	5.57	4.04	21.57	24.16
<b>Filler words</b>						
s	0.43	0.79	1.43	1.27	0.00	0.00
cl	1.14	0.90	2.14	2.54	1.29	1.25
<b>Total</b>	1.57	1.13	3.57	3.55	1.29	1.25
<b>Whole-word repetitions</b>						
s	0.57	0.79	1.14	1.07	2.29	1.70
cl	0.86	1.57	2.71	1.98	3.43	3.10
<b>Total</b>	1.43	1.62	3.86	2.67	5.71	3.77
<b>Part-word repetitions</b>						
s	0.29	0.49	0.71	0.49	0.86	1.46
cl	0.14	0.38	1.00	1.53	2.14	1.57
<b>Total</b>	0.43	0.53	1.71	1.60	3.00	2.16
<b>Prolongations</b>						
s	1.14	1.07	0.14	0.38	6.86	9.06
cl	1.14	1.21	1.43	1.62	3.43	4.20
<b>Total</b>	2.28	2.06	1.57	1.72	10.29	13.16
<b>Broken words</b>						
s	0.14	0.38	0.00	0.00	0.00	0.00
cl	0.29	0.49	0.00	0.00	1.00	1.41

<b>Total</b>	0.43	0.79	0.00	0.00	1.00	1.41
<b>Revisions</b>						
<b>s</b>	0.57	0.53	1.14	0.90	0.14	0.38
<b>cl</b>	0.43	0.49	0.57	0.79	0.14	0.38
<b>Total</b>	1.00	1.15	1.71	1.25	0.28	0.49
<b>All disfluency clusters</b>	8.00	6.08	11.57	8.00	28.43	28.02
<b>All singletons</b>	8.86	3.53	6.43	2.51	14.71	9.23
<b>All disfluencies</b>	16.86	8.67	18.00	9.56	43.14	27.29

There was a significant difference between PWC and PWS in the frequency of singletons: PWS produced significantly more singleton disfluencies.

Table 3: Results of the statistical analysis (s = occurring as singletons, cl = occurring in clusters)

	Typical & PWC		Typical & PWS		PWC & PWS	
	<i>Z</i>	<i>p</i>	<i>Z</i>	<i>p</i>	<i>Z</i>	<i>p</i>
<b>Fillers</b>						
<b>s</b>	-2.578	0.010	-	-	-	-
<b>Filler words</b>						
<b>s</b>	-	-	-	-	-2.614	0.009
<b>Whole-word repetitions</b>						
<b>cl</b>	-	-	-2.746	0.006	-	-
<b>Total</b>	-	-	-2.323	0.020	-	-
<b>Part-word repetitions</b>						
<b>Total</b>	-2.047	0.041	-2.849	0.004	-	-
<b>Prolongations</b>						
<b>s</b>	-2.152	0.031	-	-	-2.267	0.023
<b>All singletons</b>	-	-	-	-	-2.380	0.017
<b>All disfluencies</b>	-	-	-1.983	0.047	-2.366	0.018

Typical speakers produced significantly more fillers and prolongations as singletons than PWC. However, the total of part-word repetitions was signifi-

cantly more frequent in PWC than in typical speakers. PWS produced significantly more whole-word repetitions (both in clusters and total) and total part-word repetitions than typical speakers. There was significant difference between PWC and PWS in the frequency of filler words as singletons and prolongations as singletons. The former one did not appear at all, the latter one was much more frequent in the speech of PWS.

The frequency of complex disfluencies were analysed, too (Table 4). Typical speakers produced only 4 of this type. This means that three typical speakers did not produce complex disfluencies at all, and four typical speakers produced only one of them. PWC produced altogether 10 complex disfluencies. In this group, there were also three speakers who did not produce this type of disfluency. However, there was another speaker who produced 4 complex disfluencies in the 200 syllables. PWS produced 28 complex disfluencies. In this group, there was only one speaker whose speech did not contain this phenomenon. PWS with the most severe stuttering produced 10 complex disfluencies in the 200 syllables.

Utterance (1) is an example for complex disfluency containing three elements: two fillers and one part-word repetition occur in this one. Utterance (2) is an example for complex disfluency containing many elements: eight fillers, two whole-word repetitions and one prolongation occur. (FIL = filler, SIL = silent pause)

(1) **FIL SIL** sz **SIL FIL** szörnyű helytelen volt

‘it was FIL SIL a SIL FIL awfully incorrect’

(2) **eljutottam oda hogy SIL hogy FIL SIL FIL SIL hogy hogy FIL SIL egy FIL SIL bizonyos FIL SIL FIL SIL FIL SIL FIL társaságban**

‘I got to a point that SIL that FIL SIL FIL SIL that that FIL SIL a FIL SIL cerrrtain FIL SIL FIL SIL DFIL SIL FIL in the company of...’

According to the Pearson-correlation, considering every speaker, there was strong significant positive correlation between the frequency of total disfluencies and the frequency of disfluency clusters ( $r = 0.948$ ;  $p < 0.001$ ), between the frequency of total disfluencies and the frequency of complex disfluencies ( $r = 0.930$ ;  $p < 0.001$ ), and between the frequency of disfluency clusters and the frequency of complex disfluencies ( $r = 0.942$ ;  $p < 0.001$ ).

Table 4: Frequency of complex disfluencies in typical, cluttered and stuttered speech in 200 syllables

	Mean	SD	Minimum	Maximum
<b>Typical speakers</b>	0.57	0.53	0	1
<b>PWC</b>	1.43	1.62	0	4
<b>PWS</b>	4.00	3.27	0	10

### Discussion and conclusion

This paper analysed the frequency of disfluencies in the speech of PWC, PWS and typical speakers depending on whether they occurred as singletons or in clusters. Results confirmed the fact from previous studies, namely that there are not any differences between PWC and typical speakers in the frequency of the total of disfluencies, but there is significant difference between PWS and typical speakers. There had not been any data regarding the comparison between PWC and PWS. According to the results, there is significant difference between the two groups in the total of the disfluencies. Results – on the one hand – might contribute to the differential diagnosis of cluttering and stuttering, and might add to a more precise definition of cluttering, on the other hand.

Similarly to these results, Myers et al. (2012) did not find any differences between PWC and typical speakers in the frequency of the total disfluencies. They explain their results by the fact that cluttering is a multidimensional disorder and one factor (the frequency of disfluencies) is not enough for a proper diagnosis. In case of this study, the explanation is the same. However, stuttering might be well-distinguished from typical speech and cluttering, based on the frequency of disfluencies. This is the main symptom of stuttering, although there are large individual differences among PWS.

As regards to the types of disfluencies, this current paper presents different results than the previous studies did. Unlike the results of Myers et al. (2012), PWC and typical speakers of this study differed in the frequency of fillers as singletons, in the total of part-word repetitions and in prolongations as singletons. Between PWS and typical speakers, there were significant differences in the frequency of whole-word repetitions in clusters, in the total of whole-word repetitions, and in the total of part-word repetitions. Between PWC and PWS, there were significant differences in the frequency of the total of singletons, in filler words in singletons, and prolongations in singletons. The reason for these above mentioned differences from previous studies might be the small number of participants and the more dominant role of individual characteristic features.

The analysis had some limitations. One of them is the relative small number of participants. It is very difficult to find PWC with pure cluttering.



(There are other papers analysing speech of a few speakers, too; e.g. Bakker et al. 2011.) However, only 7 speakers per group are not enough for generalizing the results, because their data show individual characteristics. In addition, there were PWC and PWS with fluency disorders of different severity. For this reason, it might happen that mean values showed big differences between the groups, but according to the statistical analysis there were not any significant differences between them. There were speakers in both PWC and PWS groups who showed mild symptoms and others, who showed severe symptoms. This is why the results vary between 0 and very frequent, as regards to disfluency types.

Finally, we must not disregard the fact that recordings were made in one speech task as a structured interview. It is well known that PWC speak more fluently and show less symptoms in laboratory conditions. The stress of being recorded affects PWS exactly the opposite way as it does PWC. It is more difficult for PWS to speak. In their speech, there will be more frequent occurrences of the symptoms of stuttering.

Despite its limitations, this analysis might provide useful further information to the definition of cluttering, and to the differential diagnosis of cluttering and stuttering, and may also show a direction for future research.

### References

- Andrade, Claudia Regina Furquim de, Cervone, Luciana Maluf, & Sassi, Fernanda Chiarion 2003. Relationship between the stuttering severity index and speech rate. *Sao Paulo Medical Journal* 121/2. 81–84.
- Bakker, Klaas – Myers, Florence L. – Raphael, Lawrence J. – St. Louis, Kenneth O. 2011. A preliminary comparison of speech rate, self-evaluation, and disfluency of people who speak exceptionally fast, clutter, or speak normally. In Ward, David – Scott, Kathleen Scaler (eds.): *Cluttering. A Handbook of Research, Intervention and Education*. Psychology Press, Hove and New York. 45–65.
- Bóna, Judit 2008. A nyújtás sajátosságai a dadogó és az ép beszédben [Characteristics of prolongation in stuttered and typical speech]. *Beszédkutatás* 2015. 148–156.
- Bóna, Judit 2012. Linguistic-phonetic characteristics of cluttering across different speaking styles: a pilot study from Hungarian. *Poznań Studies in Contemporary Linguistics* 48. 203–222.
- Bóna, Judit 2016. Characteristics of pausing in normal, fast and cluttered speech. *Clinical Linguistics & Phonetics* 30/11. 888–898.
- Campbell, June Haerle – Hill, Diane G. 1994. *Systematic disfluency analysis*. Northwestern University, Evanston, IL.
- Feketéné Gacsó Mária – Mácsainé Hajós Katalin 2007. A dadogás és a hadarás vizsgálata. In Juhász Ágnes (ed.): *Logopédiai vizsgálatok kézikönyve*. [Handbook of Logopedic Examinations] Logopédiai Kiadó, Budapest. Annex 20.
- Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In Hardcastle, William J. – Laver, John – Gibbon, Fiona E. (eds.): *The Handbook of Phonetic Sciences*. Blackwell Publishing Ltd., Oxford, UK. 521–602.

- Gósy, Mária 2012a. Az artikuláció leállása a spontán beszédben. In Navracsics Judit – Szabó Dániel (szerk.): *Mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III.* Tinta Könyvkiadó, Budapest, 149–161.
- Gósy, Mária 2012b. BEA – A multifunctional Hungarian spoken language database. *The Phonetician* 105-106. 50–61.
- Gósy Mária 2012c. Sorozatmegakadások mintázata a spontán beszédben. [Series of disfluencies in spontaneous speech] *Beszéd kutatás* 2012. 107–131.
- Gósy, Mária – Bóna, Judit 2011. A case study on the effect of discourse type on fluency level in stuttering. *The Phonetician* 103-104/1-2. 57–76.
- Heeman, Peter A. – McMillin, Andy – Yaruss, J. Scott 2006. An annotation scheme for complex disfluencies. In *INTERSPEECH* 2006. <http://cslu.ohsu.edu/~heeman/papers/2006-icslp.pdf>
- Horváth, Viktória 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57/2. 288–306.
- Hubbard, Carol P. – Yairi, Ehud 1988. Clustering of disfluencies in the speech of stuttering and nonstuttering preschool children. *Journal of Speech and Hearing Research* 31/2. 228–233.
- Kántor Gyöngyi 2015. Felnőtt dadogók beszédprodukciójának műszeres elemzése és olvasásuk vizsgálata fMRI-vel. [Analysis of speech production of persons with stuttering (PWS) and reading of PWS with fMRI] Doktori értekezés. Pécsi Tudományegyetem.
- LaSalle, Lisa R., – Conture, Edward G. 1995. Disfluency clusters of children who stutter: Relation of stutterings to self-repairs. *Journal of Speech, Language, and Hearing Research* 38/5. 965–977.
- Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. The MIT Press, Cambridge (Massachusetts)–London (England).
- Lickley, Robin J. 2015. Fluency and disfluency. In Redford, Melissa A. (ed.): *The handbook of speech production*. John Wiley & Sons, Inc, Hoboken, NJ. 445–474.
- Myers, Florence L. – St. Louis, Kenneth O. 1996. Two youths who clutter, but is that the only similarity? *Journal of Fluency Disorders* 21/3-4. 297–304.
- Myers, Florence L. – Bakker, Klaas – St. Louis, Kenneth O. – Raphael, Lawrence J. 2012. Disfluencies in cluttered speech. *Journal of Fluency Disorders* 37/1. 9–19.
- Oliveira, Cristiane Moço Canhetti de – Broglio, Gabriela Aparecida Fabbri – Bernardes, Ana Paula Lázarin – Capellini, Simone Aparecida 2013. Relationship between speech rate and speech disruption in cluttering. *CoDAS* 25/1. 59–63.
- Postma, Albert – Kolk, Herman 1993. The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech, Language, and Hearing Research* 36/3. 472–487.
- Robb, Michael P. – Sargent, Ainsley – O'Beirne, Greg A. 2009. Characteristics of disfluency clusters in adults who stutter. *Logopedics Phoniatrics Vocology* 34/1. 36–42.
- Roberts, Patricia M. – Meltzer, Ann – Wilding, Joanne 2009. Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of communication disorders* 42/6. 414–427.
- Shriberg, Elizabeth Ellen 1994. *Preliminaries to a theory of speech disfluencies* (Doctoral dissertation, University of California, Berkeley).

- St. Louis, Kenneth O. – Hinzman, Audrey R. – Hull, Forrest M. 1985. Studies of cluttering: Disfluency and language measures in young possible clutterers and stutterers. *Journal of Fluency Disorders* 10/3. 151–172.
- St. Louis, Kenneth O. – Schulte, Katrin 2011. Defining cluttering: the lowest common denominator. In Ward, David – Scott, Kathleen Scaler (eds.): *Cluttering. A Handbook of Research, Intervention and Education*. Psychology Press, Hove and New York. 233–253.
- Tetnowski, John A. – Scott, Kathleen Scaler 2010. Fluency and Fluency Disorders. In Damico, Jack S. – Müller, Nicole – Ball, Martin J. (eds.): *The Handbook of Language and Speech Disorders*. Wiley-Blackwell, Oxford, UK. 431–454.
- van Zaalén-op't Hof, Yvonne – Wijnen, Frank – Dejonckere, Philippe H. 2009a. Differential diagnostic characteristics between cluttering and stuttering – Part one. *Journal of Fluency Disorders* 34/3. 137–154.
- van Zaalén-op't Hof, Yvonne – Wijnen, Frank – Dejonckere, Philippe H. 2009b. Language planning disturbances in children who clutter or have learning disabilities. *International Journal of Speech-language Pathology* 11/6. 496–508.
- Ward, David 2006. *Stuttering and cluttering: frameworks for understanding and treatment*. Psychology Press, West Sussex, UK.
- Vasić, Nada – Wijnen, Frank 2001. Stuttering and speech monitoring. In *ISCA Tutorial and Research Workshop (ITRW) on Disfluency in Spontaneous Speech*. 13–16.
- Wingate, Marcel E. 1964. A standard definition of stuttering. *Journal of Speech and Hearing Disorders* 29/4. 484–489.

### Köszönetnyilvánítás

The author wishes to thank Mária Gósy, Zsófia Koren-Dienes, Ágnes Jordanidisz and Heidi Meiners for their help in preparing this paper.



Supported through the New National Excellence Program of the Ministry of Human Capacities.

### Megakadások és megakadásklaszterek hadaró, dadogó és tipikus beszélők beszédében

A spontán beszédben megjelenő megakadások állhatnak önmagukban vagy klasztert alkothatnak más megakadásokkal. Megakadásklaszternek nevezzük azokat az eseteket, amikor ugyanazon szón vagy két egymást követő szón több megakadás fordul elő. A jelen tanulmányban azt vizsgálom, hogy milyen különbségek vannak a magukban megjelenő és a klaszterben megjelenő megakadások gyakoriságában és típusaiban a hadaró, a dadogó és a tipikus beszélők között. Elemeztem azt is, hogy milyen gyakorisággal jelennek meg

a komplex megakadásklaszterek, amelyek több mint 2 megakadás együttes megjelenéséből állnak.

Az eredmények azt mutatják, hogy a legnagyobb különbséget a megakadások típus szerinti gyakoriságában a tipikus beszélők és a dadogók között találjuk. A tipikus beszélők és a hadarók között nem volt különbség a megakadások gyakoriságában, csak egyes típusaiban. A dadogók és a hadarók szignifikánsan több komplex megakadást produkáltak, mint a tipikus beszélők.

## KÍSÉRLETEK A WAVENET MÓDSZER ALKALMAZÁSÁRA MAGYAR BESZÉDSZINTÉZISHEZ

Zainkó Csaba – Gyires-Tóth Bálint – Németh Géza –  
Olaszy Gábor

Budapesti Műszaki és Gazdaságtudományi Egyetem,  
Távközlési és Médiainformatikai Tanszék

### Bevezetés

A gépi beszédkeltés fejlődését a tudományos eredmények mellett alapvetően meghatározzák az aktuálisan elérhető gépi számítási és tárolási kapacitások is. A formáns alapú beszédszintézis a digitális szűrőkön és azok vezérlésén alapult, kihasználva az adott korban elérhető eszközök lehetőségeit (Klatt 1987; Olaszy 1985). A hullámforma összefüzetes eljárások a 90-es évek elején kezdtek elterjedni, amikor már lehetőség volt a szintézishez szükséges beszédhangminták időtartománybeli reprezentációjának tárolására és futás idejű feldolgozására (Olaszy et al. 2000). Később a háttértárak és memória-kapacitások növekedése lehetőséget nyitott a korpusz alapú beszédszintézisnek, amely akár több gigabájt mennyiségű előre rögzített hangfelvételtől intelligens eljárással válogatja össze a szintetizáláshoz szükséges hullámforma elemeket. A korpusz alapú beszédszintetizátorokhoz (Fék et al. 2006) szükséges nagy mennyiségű felvételek lehetővé tették, hogy az addig döntően szabály alapú gépi megoldások után fejlődésnek induljanak a statisztikai elveken működő megoldások (Zen et al. 2007). A rejtett Markov-modell (*Hidden Markov Model, HMM*) alapú beszédszintetizátorok már nem hullámforma hangszelletekből építik fel a szintetizált beszédet, hanem gépi tanulás útján a beszédjel spektrális tartalmából meghatározott statisztikai paramétersorozatok segítségével beszédkódolót vezérelnek (Zen et al. 2009; Tóth et al. 2012). A fenti eljárásokról részletes információk olvashatók Németh–Olaszy (2010) összefoglaló munkájában. A WaveNet is a gépi tanulásra alapozott technológia, azonban közvetlenül a hullámformára alkalmazva. Ennek a módszernek a megvalósítását az új tudományos eredményeken felül segítette az ugrásszerűen növekvő gépi számítási kapacitás (Fan et al. 2014; Wu et al. 2015; Zen et al. 2013). A módszer azért ígéretes, mert helyettesíti a HMM-nél alkalmazott beszédkódolót, ami az átalakítás során torzítást okozhat.

Magyar nyelvű WaveNet beszédszintetizátor kutatásának megindításával azt cél tűztük ki, hogy beszédminőség tekintetében megközelítsük a termé-

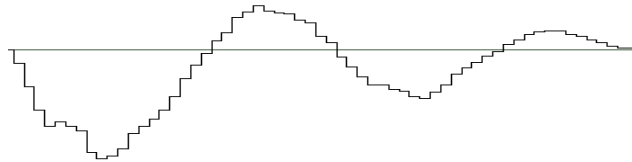
Zainkó Csaba – Gyires-Tóth Bálint – Németh Géza – Olaszy Gábor 2018. Kísérletek a Wavenet módszer alkalmazására a magyar beszédszintézishez. *Beszédkutató* 2018. 236–249.

szetes ejtésű beszédet. Jelen cikkünkben azt a hipotézisünket vizsgáljuk, hogy a korábbi WaveNet kísérleteinknél (Zainkó et al. 2017b) alkalmazott modellek finomításával javítani tudjuk-e a beszéd minőségét.

### Módszertan

A WaveNet eljárás konvolúciós neurális hálózatok (*Convolutional Neural Network*, *CNN*) segítségével végzi a gépi tanulást (Le Cun–Bengio 1995; Oord et al. 2016b). A neurális hálózatok konvolúciós rétegei igen hatékonyan képesek a jellemzőket megtanulni (*feature learning*). Ez annyit jelent, hogy magukból a nyers adatokból tanulja meg a rendszer, hogy milyen absztrakció írja le legjobban azokat. Ilyen jellegű magyar nyelvű WaveNet kísérleteket 2016 novemberében kezdtük el (Zainkó et al. 2017a).

A tanulás a digitalizált beszéd hullámformájának egyes kvantált mintái alapján történik (1. ábra), tehát közvetlenül a fizikai jelből. A kimeneten pedig a szintetizálendő hullámformát reprezentáló minták sorozata jelenik meg. A WaveNet hálózat kialakítását Oord et al. (2016c) képekre alkalmazott PixelCNN neurális hálózat architektúrája és Jozefowicz et al. (2016) szövegre alkalmazott megoldása inspirálták. A WaveNet esetén azt feltételezték, hogy ha a PixelRNN hálózat képes 64x64 pixeles képeket modellezni, akkor az audiojelek finom struktúráját is lehetséges egy hasonló módszerrel kezelni.



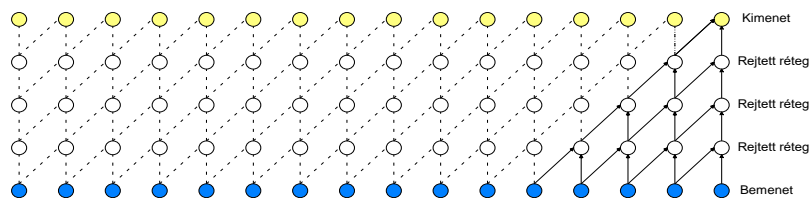
1. ábra

Egy beszédhangrészlet hullámformáját leíró mintavételi pontok az idő-amplitúdó síkon. Ezek helyét és amplitúdó számértékét használja a WaveNet a gépi tanuláshoz

A WaveNet rendszer kialakításához a PixelCNN-nél (Oord et al. 2016b) használt modell felépítését vették alapul, ahol egy képpont generálását a korábbi képpontoktól függő feltételes valószínűségek segítségével adták meg. Az  $\mathbf{x}=\{x_1, \dots, x_T\}$  hullámformához tartozó feltételes valószínűségeket az (1) képlet adja meg. Minden  $x_t$  minta függ a korábbi időpillanatok mintáitól.

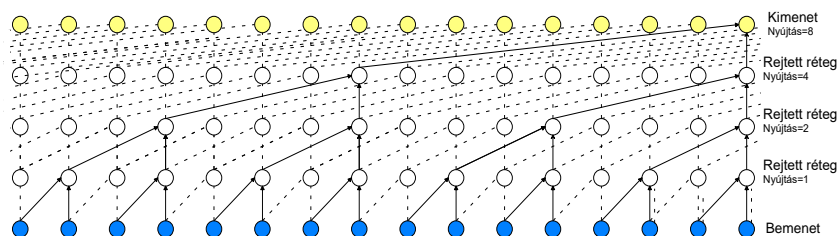
$$p(\mathbf{x}) = \prod_{t=1}^T p(x_t \mid x_1, \dots, x_{t-1}) \quad (1)$$

Komplex adatstruktúrák esetén a konvolúciós hálózatban ahhoz, hogy nagyszámú korábbi mintát figyelembe tudjunk venni, alapvetően nagyszámú rejtett réteg vagy nagyméretű szűrők alkalmazása szükséges (2. ábra). Ezeknek viszont óriásira nőhet a számítási költsége mind a tanítás, mind a generálás során, ezért az ún. nyújtott konvolúció (*dilated convolution*) architektúrát alkalmazták. Ennek lényege, hogy a rétegek nagyobb részénél, nem az előző időpillanat mintájához tartozó pontokat vonják be a konvolúcióba, hanem paraméterként megadható módon elemeket hagynak ki (3. ábra). A 2. és a 3. ábrán is 5 rétegű hálózatot látunk. Míg az első hálózat esetében a kimenet az azt megelőző 5 mintától függ, addig a nyújtott konvolúció esetén nagyságrendileg azonos számítás mellett, 16 mintától.



2. ábra

Példa egy 5 rétegű konvolúciós hálózati megoldásra  
(Oord et al. 2016a alapján)



3. ábra

Példa egy 5 rétegű nyújtott konvolúciós hálózati megoldásra  
(Oord et al. 2016a alapján)

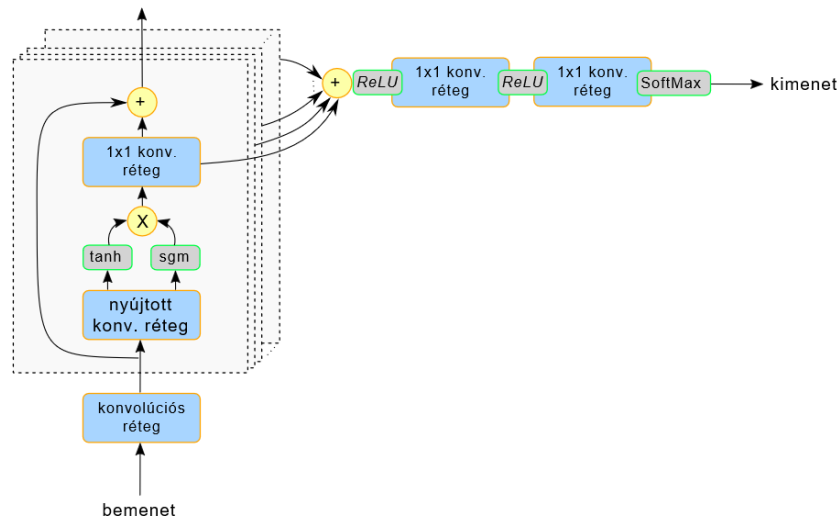
Az audiojelek előállítására tekinthetünk regressziós feladatként, de a digitális jelfeldolgozáshoz és átvitelhez széles körben használt logaritmikus kódolás segítségével osztályozási feladattá lehet átalakítani a problémát. A WaveNet esetében az ITU-T (1988)  $\mu$ -law kódolását használták. A beszédfeldolgozásban tipikusan használt 16 bites lineáris PCM jelet – amely 65536 különböző kvantálási szinttel rendelkezik – átalakítják egy 256 logaritmikus kvantálási szinttel rendelkező  $\mu$ -law kódolásba. Ezt a 256 szintű reprezentá-

ciót utána „one-hot” kódolással adják a hálózat bemenetére, ahol a 256 bemenet közül mindig csak egy tartalmaz nullától eltérő értéket.

A WaveNet esetében ún. kapuzott aktivációs (*gated activation*) egységeket alkalmaznak két aktivációs függvény szorzataként:

$$\mathbf{z} = \tanh(W_{f,k} * \mathbf{x}) \odot \sigma(W_{g,k} * \mathbf{x}) \quad (2)$$

A  $*$  jelöli a konvolúciót, az  $\mathbf{x}$  a réteg bemenet, a  $W_{f,k}$  a  $k$ -dik réteghez tartozó szűrő súlymátrixa, a  $W_{g,k}$  pedig a kapu súlymátrixa. A  $\sigma$  a szigmoid függvényt jelöli, a kör pedig az elemenkénti szorzást.



4. ábra

Rétegek kapcsolata egymáshoz és a kimenethez a WaveNet esetében  
(Oord et al. 2016c alapján)

A 4. ábrán látható, hogy a kimenetekhez minden rétegből kivezetjük az adatokat, és azok összegzése és 1x1-es konvolúciója után egy softmax függvény adja meg a kimeneti kvantált amplitúdó osztályt.

A WaveNet hálózat ebben a formában csak feltétel nélküli hullámforma generálásra alkalmas, tehát nem irányítható. Ahhoz, hogy generáláskor szabályozni tudjuk paraméterek segítségével a generálási folyamatot – amire beszéd szintézisnél szükség van (pl. bemeneti szöveg, prosódiai jellemzők) – a hálózat rétegeibe vezetjük be a szükséges információkat. A magyar WaveNet kísérletek elvégzéséhez GPU (Grafikus processzor) alapú mély tanuló keretrendszert használtunk. Az adatbázisok előfeldolgozása C++ nyelven történt, a



tanítást és a generálást pedig Python alapú TensorFlow (v1.0.0) keretrendszerrel végeztük. A keretrendszer 5.1-es cudnn-t és 7.5-ös CUDA drivert használt. A tanítást GeForce GTX TITAN X-en végeztük.

A különböző tanításokhoz és kísérletekhez az alábbi beszédatadátbázisokat használtuk:

FONETIKA korpusz (Olaszy 2013): 10 beszélővel felolvasott 2000 mondatos párhuzamos beszédatadátbázis, amelyben a hangkapcsolatok fonetikailag kiegyenlítettek. Az adatbázis összesen  $10 \times 2000 = 20.000$  mondatot tartalmaz.

MK\_MÁV egybeszélős felolvasott beszédkorpusz, amely a MÁV állomások hangos utastájékoztató rendszeréhez optimalizált mondatokból áll (Zainkó et al. 2015). A hangfelvételek stúdióban készültek professzionális rádióbemondó közreműködésével. A korpusz 3225 annotált mondatot tartalmaz.

A hibamérés fontos eleme a WaveNet modell tanításának. A neurális hálózatok tanítása során a túltanítás elkerülésére a leállási feltételt gyakran a hibafüggvény értékének alakulásához kötjük. A hibafüggvény azt adja meg, hogy mekkora a különbség a tanító adat és a modell által aktuálisan előre jelzett (generált) minta között. Például, ha adott tanítási cikluson keresztül nem csökken a hiba (vagy elkezd növekedni) egy tanító adatoktól elkülönített, ún. validációs halmazon, akkor leállítjuk a tanítást. A validációs halmazt úgy alakítjuk ki, hogy az eredeti tanító adatok egy részét elkülönítjük, azok nem vesznek részt a tanításban. A WaveNet hálózat tanítása két szempontból is speciális. Egyrészt a generált hangmintákat megvizsgálva jellemző, hogy nem minden esetben a legkisebb hibát produkáló hálózatok adják a legjobb szubjektív értékelést a tesztelőknél. A másik tényező az, hogy a validációs halmaz elemeivel való teljes összehasonlításhoz a mondatokat le kell generálni. Ez nagyon időigényes folyamat, ezért a hatékony tanítás érdekében maradtunk a hiba egyszerű mérésénél, tehát a hálózat hibáját vizsgáltuk a validációs halmazon anélkül, hogy generálást végeztünk volna. A hiba mérésekor tehát csak mindig a következő előrejelzett minta értékét hasonlítjuk össze a valódi minta értékével. Ez gyorsabb, de kevésbé pontos.

A WaveNet eljárás során egymástól független tanítási és generálási (szintézis) fázisokat különböztetünk meg. A tanítási fázis alatt a tanító adathalmaz segítségével a modellt úgy próbáljuk meg módosítani, hogy a validációs halmazon mért hiba folyamatosan csökkenjen. A tanítás során a kísérletekkel határozzuk meg, hogy milyen méretű hálózat alkalmas jobb beszédminőség elérésére. Szintén a kísérletekkel vizsgáljuk, hogy milyen vezérlőparaméterek használata esetén tudjuk az emberihez minél közelebbi hangminőséget elérni. A generálási (szintézis) fázisban a modellt kismértékű zajjal indítjuk el, majd a vezérlő paramétereket figyelembe véve mintánként generálja a beszéd hullámformáját. A generálás esetében mindig az előzőleg generált mintákat használja a hálózat. Mivel a következő minta generálásához szükséges az

előzőleg generált minta, ezért nem tudjuk a GPU-k párhuzamos számítását kihasználni és egyszerre több mintát előállítani.

A szintézis gyorsítható, Le Paine (2016) Fast-WaveNet módszerével. A forward lépéseknél olyan számításokat végzünk el minden egyes lépésnél, amit már korábban egyszer kiszámoltunk. La Paine rámutatott, hogy a rész-eredmények eltárolásával a generálás  $O(2^L)$ -ről  $O(L)$ -re gyorsítható, ahol  $L$  a rejtett rétegek száma. A saját méréseink először nem támasztották alá ezt a gyorsulást. A generálás folyamatát elemezve megállapítottuk, hogy a Fast-WaveNet esetében a numerikus számítás mennyisége annyira lecsökken, hogy a futási idő legnagyobb részét a GPU-ra történő adatátvitel és a számítások után az eredmények memóriába való visszaolvasása adta.

### Kísérletek

Az első kísérletünkben azt vizsgáltuk, hogy a különböző vezérlő paramétereket hogyan tudjuk felhasználni a modell vezérlésére, illetve milyen hatással vannak közvetlenül a hibára, közvetve pedig a generált beszéd minőségére. A tanító adatbázis összetétele és a hálózat mérete változtatásának hatását a második kísérletben vizsgáltuk. Ebben három különböző modellt készítettünk és ezekkel szintetizáltunk tesztmondatokat. Ahhoz, hogy a fejlődést nyomon tudjuk követni, korábbi (Zainkó et al. 2017b) modellek segítségével is elkészítettük ugyanazokat a mondatokat, majd meghallgatásos teszttel értékeltük őket.

### A mély rétegek szabályozása

A WaveNet hálózat nem csak a bemeneti rétegen keresztül vezérelhető, hanem minden rétegben módosíthatjuk a szűrő és a kapu súlyok hatását (Oord et al. 2016a). Az (1)-es képletet módosítva a feltételes eloszlásunk a következő formába írható át:

$$p(\mathbf{x} | \mathbf{h}) = \prod_{t=1}^T p(x_t | x_1, \dots, x_{t-1}, \mathbf{h}) \quad (3)$$

A plusz bemeneteket  $\mathbf{h}$ -val jelöltük. A  $\mathbf{h}$  lehet egy globális paraméter, amely hosszú időn keresztül állandó, például a beszélő azonosítója. Amennyiben egy  $y=f(h)$  függvénnyel a bemeneti mintákhoz illesztett paraméterlistát generálunk (például hangkódok vagy egyéb nyelvi jellemzők), akkor a konvolúciós egységekben lévő aktivációt - (2)-es képlet - a következőképpen módosíthatjuk (ahol  $V_{f,k} * \mathbf{y}$  és  $V_{g,k} * \mathbf{y}$  egy-egy 1x1-es konvolúció):

$$\mathbf{z} = \tanh(W_{f,k} * \mathbf{x} + V_{f,k} * \mathbf{y}) \odot \sigma(W_{g,k} * \mathbf{x} + V_{g,k} * \mathbf{y}) \quad (4)$$

A kísérleteinkben a -2, +2 beszédhangos környezettel kibővített beszédhangot (5 hangnyi területet) kódoltuk, illetve az alaphangfrekvencia logaritmusát.

Ezek mellett további hangkörnyezetet leíró jellemzőket használtunk fel. Például azt, hogy a középső hang zöngés, vagy magánhangzó-e, illetve azt, hogy mennyi a hang időtartama. Ezek mellett prozódiai információkkal kapcsolatos paramétereket is felhasználunk, mint például a vizsgált hangot tartalmazó szó pozíciója a mondatban vagy azt, hogy a szó a mondat melyik prozódiai egységében található (kezdő, közbenső, utolsó, vagy csak egy ilyen egységből áll a mondat).

Több-beszélős adatbázis esetében nem használtuk a *h* globális paraméterezés lehetőségét, a beszélő azonosítóját is az *y* bemenetre konvertáltuk át, illetve többnyelvű modell esetén a beszéd nyelvi azonosítóját is hasonlóan kezeltük.

#### **A paraméterek szerepe a generálásban**

A vezérlő információk nem egyforma mértékben hatnak a kimeneti hálómforma minőségére. Az első kísérletben azt vizsgáltuk, hogy ha egy paramétert nullázunk (hatását kiiktatjuk), akkor mennyivel növekszik meg a hiba a validációs halmazon. Amennyiben a hiba alig változik, az azt jelenti, hogy kevésbé fontos a hálózat számára ez a paraméter, ha pedig az adott paraméter elhagyása jelentősen növeli a hibát, akkor az fontosabb a vezérlés szempontjából.

#### **A modellek tanítása**

A második kísérletben azt vizsgáltuk, hogy miként lehet egy sokbeszélős modellt felhasználni egyetlen beszélő jó minőségű szintetizálásához. A korábbi WaveNet kísérletek rámutattak, hogy több beszélő adatainak párhuzamos használata segíti, hogy a rendszer ne az adott beszélő egyéni tulajdonságait, hanem a több beszélő közös tulajdonságait tanulja meg. Ez azt jelenti, hogy a rendszer képes az adott nyelv jellemzőinek egy részét megtanulni. Lehetőség van arra, hogy az elkészült modellt úgy vezéreljük, hogy egy adott beszélő hangját generálja, de a beszélők egy adott csoportjának együttes beszédét is generálni tudja, tehát egy átlaghangon állíthat elő beszédet. Ez lehetőséget ad arra, hogy az egyes beszélők hibáit a rendszer figyelmen kívül hagyja, jobb minőségű, a nyelvre jobban jellemző beszédet szintetizáljon.

A modell tanítása két lépésből állt. Első lépésként sok beszélő segítségével egy átlaghangot tanítottunk. A FONETIKA adatbázis 10 beszélője segítségével tanítottuk a rendszert. A mondatok és a beszélők is véletlen sorrendben követték egymást. A tanítás után a modell képes volt mind a 10 beszélő hangján beszédet előállítani vagy azok tetszőleges kombinációjával. Például lehetőség van nő vagy férfi átlaghang megszólaltatására, vagy magyar átlaghangú beszéd generálására, amelynél a beszélő neme nem megállapítható. A személyparaméterek folyamatos változtatásával személyek közötti átmenet is megvalósítható.

A HMM beszédsszintetizálási modelleknél jellemző tanítási módszer, hogy sok beszélő átlaghang modelljét adaptálják egy beszélőre. WaveNet esetében is lehetőség van erre, így a kísérleteink során a fent elkészült több-beszélős modellt tanítottuk tovább, már egyetlen hanggal úgy, hogy a továbbiakban már nem adtuk meg a modellnek, hogy melyik beszélő hangját tanulja, hanem az átlaghangot módosítja már.

A tanítási folyamat második lépésében az MK\_MÁV adatbázis egyetlen beszélőjével folytattuk a tanítást, a beszélő paramétereit átlaghangnak állítottuk be. A modell ekkor már csak ennek az adott beszélőnek a hangjellegzetességeit tanulta.

A modellek tanítása során 3 különböző rendszer készítettünk (egyszerű, közepes, nagy), amelyeknél a tanító adatok és a tanítás módja, paraméterek száma megegyezik, csak a hiper-paraméterekben van különbség. Az egyszerű modell esetében a konvolúciós szűrő mélysége 32 volt, a másik kettőnél 64. A rétegek száma az egyszerű és a közepes modell esetében 40 réteg volt, a nagy esetében pedig 80 rejtett réteg szerepelt a hálózatban.

### **Mondatok generálása**

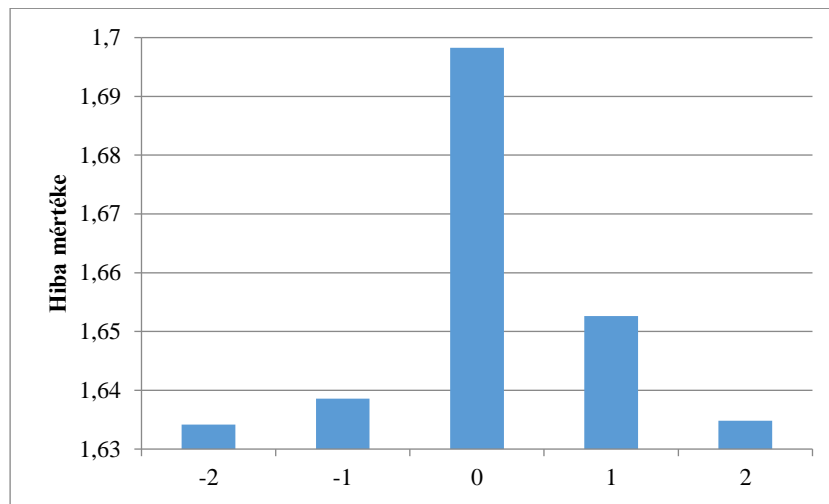
WaveNet tanítása időigényes, de a nagy mennyiségű feldolgozott adatot figyelembe véve hatékonnak tekinthető. A modell paramétereitől függően egy iterációs lépés kb. 1,4-2 mp-ig tart, amely során 100.000 mintát, 6,25 mp hanganyagot használunk fel. A tanítás és a generálás is 16kHz-es mintavételi frekvenciával történt. A mérések szerint egy minta generálása, kb. 0,25 mp-ig tart. Így nagyságrendileg 1 mp hanganyag előállítása kb. 2 óra.

A szintézis Le Paine (2016) Fast-WaveNet módszerével gyorsítva és a GPU-t kihagyva, csak CPU-n (központi processzoron) futtatva 1 mp hanganyag előállítása csak kb. 4 percet vett igénybe. A program CPU-ra történő optimalizálásával és a számítások párhuzamosításával az 1 másodperces hanganyag előállítását sikerült végül 13 mp-re leshorítani.

### **Eredmények**

#### **A paraméterek szerepe a generálásban**

Az 5. ábrán azt láthatjuk, hogy ha az adott hang vagy annak környezetéből egy hangra vonatkozó információt törölünk, mennyire nő meg a hiba mértéke. Az ábrán a hálózat átlagos hibáját ábrázoltuk.

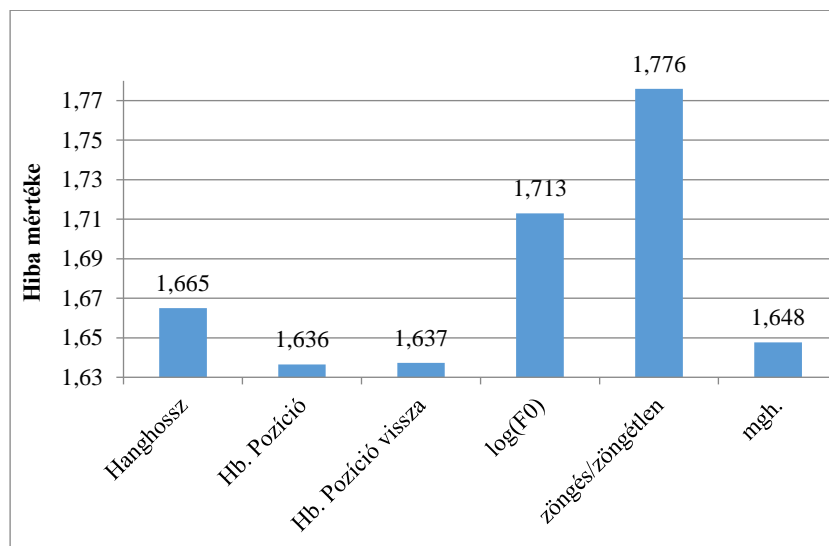


5. ábra

A hangok elhagyásának hatása a validációs hibára

A vártak megfelelően a szintetizálendő hang elhagyása okozta a legnagyobb hibát. A második legfontosabbnak az adott hang után következő hang bizonyult, majd az adott hang előtt álló következett. A legkevésbé fontos a kísérlet alapján az adott hang előtt kettővel álló beszédhang.

A további 31 különböző vezérlő paraméter hatását is vizsgáltuk. A 6. ábrán látható a 6 legfontosabb paraméter hatása a hibára. A zöngésségi paraméter elhagyása okozta a legnagyobb hibát a rendszerben. A valós érték az ábrából kilógna, értéke 1,766. Második legfontosabbnak az alaphékvencia értéke bizonyult, majd az adott hang hossza. A negyedik legfontosabb paraméter azt adja meg, hogy magánhangzó-e az adott hang. További fontos paraméter az, hogy a hangon belül a generálás hol tart, illetve mennyi van még hátra az adott beszédhang teljes megvalósításából. Az ábrán nem megjelenített paraméterek is hatással voltak kis mértékben a hibára.



6. ábra

Az első 6 legnagyobb validációs hibára ható paraméter

Mivel a WaveNet egy hullámforma generáló eljárás és magában nem alkalmas arra, hogy szövegből beszédet előállítson, ezért természetes mondatokból indulva állítottuk elő a tesztben szereplő mondatokat. Nem szövegből generáltuk a vezérlő paramétereket, hanem a természetes ejtésű mondatokból származtattuk. Ezzel az eljárással elkerülhető az egyéb modulok (fonetikai átíró, prozódiai modul) befolyása a hangminőségre, például a prozódia helytelen modellezése. A szintetizálás során a tanító adatbázisban nem szereplő mondatok vezérlő paramétereinek felhasználásával állítottuk elő a hullámformákat. Kiválasztottunk 8 különböző mondatot, majd mind a 3 új rendszerrel és a korábbi tanulmányban (Zainkó et al. 2017b) ismertetett 2 legjobb rendszerrel előállítottuk a szintetizált változatokat.

#### Meghallgatásos teszt

A második kísérlet eredményeinek értékeléséhez meghallgatásos MUSHRA (MULTI-Stimulus test with Hidden Reference and Anchor) tesztet készítettünk ITU-R (2001). A teszt lényege, hogy az alanyok a meghallgatásos összehasonlítás során egy referencia mondathoz hasonlítják a különböző módszerrel előállított mondatokat és megadják egy 100-as skálán, hogy mennyire találják a referenciához hasonlóknak a meghallgatott mondatot. A meghallgatandó mondatok közé a referencia mondatot is elrejtettünk.

A tesztbe 2 korábbi rendszert is bevettünk, hogy a modellek fejlődését is tudjuk vizsgálni. A (Zainkó et al. 2017b) cikkben szereplő rendszerek többnyelvű modellt használnak, az egyik rendszer 2 nyelvű (magyar–angol) a másik háromnyelvű (magyar–angol–német). A tesztben csak magyar nyelvű mondatokat generáltunk. A tesztben összesen 8 mondat 5 változatát kellett értékelnie az alanyoknak. A mondatok a MK\_MÁV adatbázis tematikájához illeszkedő mondatok voltak. A tesztelt rendszerek főbb tulajdonságait az 1. táblázatban adjuk meg.

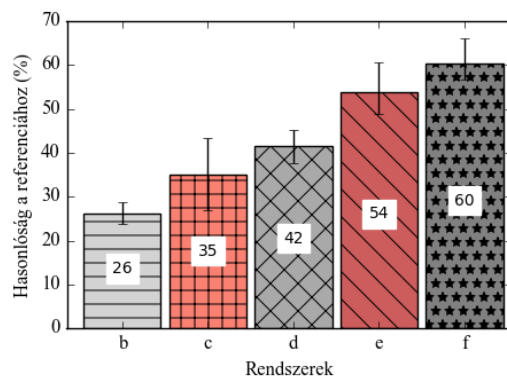
1. táblázat: A kísérletben felhasznált modellek főbb adatai

Azonosító	Rétegek száma	Szűrők mélysége	Nyelvek száma
<b>b</b>	40	64	2
<b>c</b>	40	64	3
<b>d</b>	40	32	1
<b>e</b>	40	64	1
<b>f</b>	80	64	1

A tesztet 27 felnőtt töltötte ki, 11 férfi és 16 nő. A tesztet átlagosan 12,6 perc alatt végezték el. A beszélők átlag életkora 50 (29–77) év volt.

### Eredmények

A MUSHRA tesztben előállt rangsort Mann–Whitney–Wilcoxon tesztel is összevetettük, 95%-os konfidenciaszintet használva. Ez alapján a vizsgált rendszerek mondatai szignifikánsan különböző minőségűnek bizonyultak a horgony és a WaveNet által szintetizált mondatok mindegyikéhez képest, tehát mind az 5 rendszer szignifikánsan különbözik egymástól. A legrosszabb eredményt a Zainkó et al. (2017b) cikkben leírt korábbi rendszer érte el (b modell), tehát az új modellek sikeresebbek lettek a korábbinál. Az eredmények arányosak voltak a modell méretével, minél összetettebb, nagyobb volt a modell, annál jobb eredményt ért el. Az eredmények a 6. ábrán láthatóak (a rejtett referenciát nem jelenítettük meg). A tesztalanyok a referencia mondattal pontosan megegyező mintát 92%-osra értékelték. Mivel nem volt külön negatív referencia, ezért a teszt alapján abszolút minőségi értékelést nem tudunk adni, csak a rendszereket tudjuk egymáshoz hasonlítani.



7. ábra

Az 5 vizsgált WaveNet modell hangminősége

### Összefoglalás

A magyar nyelvű WaveNet kísérleteinkkel sikerült elérni azt a beszédminőséget, amely alapján a modell felhasználható beszédsszintetizálás céljára. A beszédminőséget szignifikánsan sikerült javítanunk a korábbi rendszereinkhez képest, a modell finomítása megfelelő eredményt hozott, javult a hangminőség (f modell). A modell legnagyobb hátránya jelenleg az, hogy nem tudjuk a más beszédsszintézis technológiákhoz hasonló számítási sebességet megközelíteni. Ez jelentősen korlátozza az alkalmazási lehetőségeket. Ebben jelentős javulás várható a közeljövőben, mert a Google DeepMind kutatócsoportja már bejelentette, hogy angol és japán nyelvű megoldásaiban ezerszeres gyorsulást ért el (Oard 2017d).

A WaveNet vagy a várhatóan nagy számban megjelenő hasonló mély neurális hálózat alapú rendszerek egyértelműen jobb hangminőségű beszédet tudnak előállítani, mint a korábbi technológiák és a sebesség problémák megoldása után széles körű elterjedésük várható. Mivel ezek a technológiák már közelítik az emberi beszéd minőségét, és egyes esetekben már nem megkülönböztethető az emberi beszéd a gépi változattól, alkalmazásuknál fokozottan jogi és etikai problémák kerülhetnek előtérbe. Elképzelhető, hogy új kutatási irány lesz az, hogy a szintetizált beszéd teljesen emberi legyen, de ne hasonlítson egyetlen tanító adatbázisban szereplő beszélő hangjára sem (nyelvi átlag hang).

### Irodalom

Fan, Yuchen – Qian, Yao – Xie, Fenglong – Soong, Frank K. 2014. TTS synthesis with bidirectional LSTM based recurrent neural networks. *Interspeech*. 1964–1968.



- Fék, Márk – Pesti, Péter – Németh, Géza – Zainkó, Csaba – Olaszy, Gábor 2006. Corpus-based unit selection TTS for Hungarian. In *International Conference on Text, Speech and Dialogue* 367–373. Springer
- ITU-R. Recommendation BS.1534: Method for the subjective assessment of intermediate audio quality (2001)
- ITU-T. Recommendation G. 711. Pulse Code Modulation (PCM) of voice frequencies 1988.
- Jozefowicz, Rafal – Vinyals, Oriol – Schuster, Mike – Shazeer, Noam – Wu, Yonghui 2016. Exploring the limits of language modeling 2016 (Feb 7) arXiv preprint arXiv:1602.02410.
- Klatt, Dennis 1987. Review of text-to-speech conversion for English. *The Journal of the Acoustical Society of America* 82/3. 737–793.
- Le Cun, Yann – Bengio, Yoshua 1995. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361/10.
- Nagy, Péter – Zainkó, Csaba – Németh, Géza 2015. Synthesis of speaking styles with corpus-and HMM-based approaches. *Cognitive Infocommunications (CogInfoCom)*, 6th IEEE International Conference on. IEEE.
- Németh, Géza – Olaszy Gábor (szerk.) 2010. A magyar beszéd. Akadémiai Kiadó, Budapest.
- Olaszy Gábor 1985. A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. A számítógépes beszéd-előállítás néhány kérdése. *Nyelvtudományi Értekezések* 121. Akadémiai Kiadó, Budapest.
- Olaszy, Gábor 2013. Precíziós, párhuzamos magyar beszédatadátbázis fejlesztése és szolgáltatásai. *Beszédkutató* 2013. 261–270.
- Olaszy, Gábor – Németh, Géza – Olaszy, Péter – Kiss, Géza – Gordos, Géza 2000. "PROFIVOX - A Hungarian Professional TTS System for Telecommunications Applications", *International Journal of Speech Technology* 3. 3-4. 201–216.
- Oord, Aaron van den – Dieleman, Sander – Zen, Heiga – Simonyan, Karen – Vinyals, Oriol – Graves, Alex – Nal, Kalchbrenner – Senior, Andrew – Kavukcuoglu, Koray 2016a. WaveNet: A generative model for raw audio. arXiv preprint arXiv:1609.03499.
- Oord, Aaron van den – Nal, Kalchbrenner – Kavukcuoglu, Koray 2016b. Pixel Recurrent Neural Networks. arXiv preprint arXiv:1601.06759.
- Oord, Aaron van den – Nal, Kalchbrenner – Vinyals, Oriol – Espeholt, Lasse – Graves, Alex – Kavukcuoglu, Koray 2016c. Conditional image generation with pixelcnn decoders. arXiv preprint arXiv:1606.05328.
- Oord, Aaron van den – Li, Yazhe – Babuschkin, Igor – Simonyan, Karen – Vinyals, Oriol – Kavukcuoglu, Koray – van den Driessche, George – Lockhart, Edward – Cobo, Luis C. – Stimberg, Florian – Casagrande, Norman – Grewe, Dominik – Noury, Seb – Dieleman, Sander – Elsen, Erich – Kalchbrenner, Nal – Zen, Heiga – Graves, Alex – King, Helen – Walters, Tom – Belov, Dan – Hassabis, Demis 2017d. Parallel WaveNet: Fast High-Fidelity Speech Synthesis arXiv preprint arXiv: 1711.10433
- Le Paine, Tom 2016 (nov.10). Fast Wavenet: An efficient Wavenet generation implementation <https://github.com/tomlepaine/fast-wavenet>

- Tóth, Bálint, Pál – Németh, Géza 2009. Rejtett Markov-modell alapú szövegfeldolvasó adaptációja félig spontán magyar beszéddel. VI. Magyar Számítógépes Nyelvészeti Konferencia, Szeged, 246–256.
- Wu, Zhizheng – Valentini-Botinhao, Cassia – Watts, Oliver – King, Simon 2015. Deep neural networks employing multi-task learning and stacked bottleneck features for speech synthesis. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 4460–4464.
- Zainkó, Csaba – Bartalis, Máttyás – Németh, Géza – Olaszy, Gábor 2015. A Polyglot Domain Optimised Text-To-Speech System for Railway Station Announcements. Sixteenth Annual Conference of the International Speech Communication Association.
- Zainkó, Csaba – Tóth, Bálint Pál – Németh, Géza 2017a. Polyglot Magyar nyelvű WaveNet kísérletek, In XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2017), Szeged. 205–216.
- Zainkó, Csaba – Tóth, Bálint Pál – Németh, Géza – Olaszy, Gábor 2017b. Polyglot speech generation with WaveNet, In Technical report of Smartlab, [http://smartlab.tmit.bme.hu/downloads/pdf/zainko/Zainko\\_Polyglot\\_2017.pdf](http://smartlab.tmit.bme.hu/downloads/pdf/zainko/Zainko_Polyglot_2017.pdf) Letöltés: 2017. November 15.
- Zen, Heiga – Senior, Andrew – Schuster, Mike 2013. Statistical parametric speech synthesis using deep neural networks. In IEEE International Conference on Acoustics, Speech and Signal Processing. 7962–7966.
- Zen, Heiga – Toda, Tomoki – Nakamura, Masaru – Tokuda, Keiichi 2007. Details of the Nitech HMM-based speech synthesis system for the Blizzard Challenge 2005. IEICE transactions on information and systems 90/1. 325–333.
- Zen, Heiga – Tokuda, Keiichi – Black, Alan 2009. Statistical parametric speech synthesis. *Speech Communication*, 51/11. 1039–1064.

### Köszönetnyilvánítás

A szerzők köszönetüket fejezik az NVIDIA Corporation-nak, hogy a kutatást egy NVIDIA TITAN X GPU-val támogatták. A kutatást támogatta továbbá a DANSPLAT (EUREKA 9 944) és a VUK (AAL-2014-1-183.) projekt.

### Experiments to apply the WaveNet method for Hungarian speech synthesis

The WaveNet architecture is suitable to generate high quality speech, it was demonstrated for English by Google DeepMind. In this paper we describe our experiments of using WaveNet for Hungarian speech generation. We investigated the effects of different control parameters and compared the quality of generated speech with different hyper-parameter settings and with different Hungarian speech databases. We examined the most influential control parameters and we conducted a listening test to investigate the evaluation of Hungarian WaveNet models. Significant increase in quality of synthesized speech with larger models and with a modified approach was achieved.

## „NYELVBOTLÁS”-KORPUSZ

### 14. rész

Az első magyar valós idejű, jegyzeteléses technikával gyűjtött megakadásjelenség-korpuszt 2004 óta adjuk közre folyóiratunkban:

Beszéd kutatás 2004. 19–186.: 5139 adat;  
 Beszéd kutatás 2005. 145–173.: 761 adat;  
 Beszéd kutatás 2006. 231–247.: 388 adat;  
 Beszéd kutatás 2007. 187–198.: 244 adat;  
 Beszéd kutatás 2008. 221–239.: 444 adat;  
 Beszéd kutatás 2009. 257–267.: 176 adat;  
 Beszéd kutatás 2010. 283–291.: 179 adat;  
 Beszéd kutatás 2011. 149–165.: 429 adat;  
 Beszéd kutatás 2012. 301–313.: 298 adat;  
 Beszéd kutatás 2014. 253–276.: 604 adat;  
 Beszéd kutatás 2015. 248–256.: 193 adat;  
 Beszéd kutatás 2016. 193–214.: 545 adat;  
 Beszéd kutatás 2017. 241–255.: 377 adat.

Eddig tehát összesen 9777 megakadásjelenséget elemeztünk, ebben a részben pedig újabb 378 tételt adunk közre.

Értelemszerűen ez a korpusz a hiba típusúak gyűjteménye. A kezdetektől számos, különféle szempontú tanulmány, szakdolgozat, disszertáció készült ebben a témakörben. Mindez a tapasztalat oda vezetett, hogy a „Nyelvbtlás”-korpusz osztályozási kategóriáit 2009-től egyszerűsítsük. Az alapvető ok az volt, hogy a korpuszt használóknak nagyobb szabadságot biztosítsunk egy adott jelenség mélyreható elemzéséhez.

Az egyszerűsített osztályozás elve az, hogy minden esetben a felszínen tapasztalható jelenséget soroljuk kategóriákba, a finomelemzések az adott jelenség kutatójára várnak. A példákban a □ a néma szünetet jelöli. A fő kategóriák és meghatározásuk a következők:

#### 1. Téves szó

A megfelelő szó helyett egy másik szó megjelenése a felszínen. Ez a kategória tartalmazza a freudi elszólásokat és a malapropizmusokat is, amelyek további elemzésekkel különíthetők el.

#### 2. Grammatikai hiba

A köznyelvi normának ellentmondó morfológiai/szintaktikai struktúra.

#### 3. Kontamináció

Két nyelvi jel (szó, szerkezet) összeolvadása, vegyülése.

#### 4. Téves kezdés

Beszédhang vagy beszédhangsorozat ejtése, amely azonban nem szó.

### 5. **Nyelvem hegyén van jelenség**

A szándékolt szó jelentésének és morfológiai struktúrájának ismerete, de a fonológiai/fonetikai szerkezet kiejtésének (részleges) gátoltsága.

### 6. **Perszeveráció**

Egy kiejtett nyelvi elem szándéktalan, ismételt megjelenése a közlés egy későbbi időpontjában.

### 7. **Anticipáció**

Egy szándékozott nyelvi elem megjelenése a közlés egy korábbi időpontjában is.

### 8. **Metatézis**

Nyelvi elemek sorrendjének felcserélődése.

### 9. **Egyszerű nyelvbtlás**

A beszédtervezés artikulációs tervezési szintjén, ill. az artikulációban létrejött hiba, amely semmilyen egyéb okkal nem magyarázható. Altípusai: a) betoldás, b) helyettesítés, c) kiesés

### 10. **Több típusba sorolható jelenségek**

Azok az adatok kerülnek ebbe a kategóriákba, amelyek létrejöttében a felszíni elemzés alapján több rejtetten működő folyamat hibája is feltételezhető.

### 1. **Téves szó**

#### **Megakadás**

...az egyik a kísérletek körül □ közül...

...mert egy jó énekező, énekes...

4-es-6-ossal kell menni, de előtte még kell  
valami töményt vennem.

A □ akkor més? Vagyis jössz?

A béke legyen a híd, mely közénk áll.

A fecske is kötöző madár?

A gazdasági miniszter vagyis aaa □ a gazdasági miniszter tegnapi beszédében...

A kérdéses kérdés vagy micsoda kérdéses kutatás.

A két óra hossza, amit ott eltörtünk, eltörtünk, egynek tűnik.

A lányok sokkal kicsit sokkal közelebb vannak az átlaghoz.

A fogódzkodót megragasztva □ megragadva...

A meleget egyensúlyozza, elnézést ellensúlyozza az északi szél.

#### **Szándékolt közlés**

tömbjegyet

köztünk

költöző

A mentolom volt a gólyatáborban □ mento- rom, na!	
A város forgalmi rendjének □ vagyis for- galmi tervének elkészültével...	
...adalékmentes öőö adómentes juttatási forma...	
Akkor jöjjön le, és olvasson be, vagy le.	olvassa le (a villanyórát)
Amikor megadjuk a kezdőszót vagy na, kez- dőhangot.	
Amikor volt még beszédképzés a tévében, beszédtanítás.	
Anyukám már felhívott délnyitány.	délután
Az élet meg a rutin...	az évek meg a rutin
Az észak-korai □ vagyis koreai hadsereg ter- vez valamit.	
Az ilyen tanulmányos vagy nem, tudomá- nyos cikkeknel.	
Az olajfák hegyén következik!	olajfák
Az unokahúgom idén olvassa a Lassie haza- telefonált □ vagyis hazatért, na!	
Azt hiszi, hogy félre megyek.	félre lépek
Bármi legyen is ez hallgasd, halkítsd le, légyszi!	
Befonod a szemem?	hajam
Bizonytalanság felolvasása vagy feloldása.	
Csak megkóstoltam, milyen rumot eszel □ eszel?! Iszol, csak éhes vagyok!	
Egyszer én is fogtam már 10 kilós haját □ ha- ját □ halat!	
El kell mennem kulcsot fénymásoltatni.	másoltatni
Eleve nehéz ebben a melegben tanulni, hát még úgy, hogy nem hagy olvadni □ na, olvasni.	
Elment feljelentést venni tenni.	
Elpakolok egy táliba □ vagyis egy dobozba.	
Elvittem a kulcsmásolóba a cipőt.	cipőjavítóba hiányoltam
Én azt onnan hiányoztam.	
Én ettem már ilyet □ vagy ittam.	
Én most már lefekszem zuhanyozni.	aludni
Ez csirkenyúl?	csirkecomb
Ez is feladó azaz eladó?	
Ez milyen szőlő? Vagy szilva, bocsánat.	

Ezt az erdőből hozták?	erődből
Ezt én is el tudom foglalni.	fogadni
Ezt is átírom vagy ugrom.	
Ezután az izgató, azaz az igazgató kinevezését...	
Fáj a lábam □ na aaaaa térdem.	
Fel akartam mondani neked, hogy...	el akartam mondani
Gondolkodom a témán vagy a példán.	
Házasságtörő terem	házasságkötő terem
Hibát követünk el, azaz nem követünk el, hanem produkálunk.	
Hoztam neked csokit □ vagyis sütit.	
Hoztam tojást is! Ó tojást! Túrórudit.	
Ilyen tekintélyben, vagyis tekintetben nem.	
Inkább férfias vagy inkább nős nőies?	
Inkább részeg álomnak □ álmokfutásnak tűnt ez nekem, mint szándékos agressziónak.	
Itt nem testi bántásról van szó.	bántalmazásról
Kapcsold ki a tabletet □ vagyis a gépet!	
Kérdőjelezd meg a felső, feletted lévő embereket!	
Kértelen kelletlen	kénytelen kelletlen
Késélen táncol	pengeélen táncol
Két percet kapott a mikróban, elég magas lesz?	meleg
Kéz- és csonttörtést, vagyis lábtörtést!	
Kiettem az almát.	megettem
Kirakhatnák vízkővel .	díszkővel
Kísérletiesen hasonlít	kísértetiesen hasonlít
Kókuszos, nem koztümös film.	
Könnyebben felmerülnek, felkerülnek.	
Krattyogott □ vagyis kattogott.	
Küldd el azt a borítékot még ma!	levelet
Legyen szíves, csukja be az ajtót □ ajtóóóót □ablakot!	
Lehetek még szentelenebb? – Menjél, vagy legyél.	
Lelkileg is tör.	gyötör
Magam alatt ásom a fát.	vágom
Majd én gondolkodom rólad!	gondoskodom
Majd három vagy négy órát □ háromnegyed órára bezárkózott.	
Már is megyek zoknikám!	Zsoltikám

Marokkó is szuvenír ország már.	szuverén
Meghívás keringőre	felhívás keringőre
Megírtad a füzeted? Vagyis a házit? Ne piszkáld már a füzeted!	
Merre van a módosítószó?	mosdó
Mész tornázni?	fürödni
Mi is itt lettünk, na, itt jöttünk be.	
Mi újság a török □ torokfájásoddal?	
Nagyapámék még az Arany János utcai parókán laktak a szolgálati ideje alatt.	parókián
Ne beszélj hozzám! Vagyis vissza.	Ne beszélj vissza!
Néhány honlap ír vele ööö él vele.	
Nem dallamban ööö szólamban énekelte.	
Nem egy megfelelő □ meglepő dolog.	
Nem fogom rá használni, sajnálni, nem fogom rá sajnálni a pénzt.	
Nem hoztam lapátot.	kabátot
Nem szabad Marika!	Marcika
Nem találok a bérletem □ akarom mondani a kulcsom.	
Nem tudtam beazonosítani, de patásujjú párosujjú patás lehetett.	
Nem túl gyakoriság.	gyakori
Nyisd szét az ajtót.	nyisd ki
Ő csak hétfőn meg szerdán vagy csak benn van csak benn.	
Ő is muszlin? Akarom mondani, muszlim.	
Ő meg látom, nőül □ vagyis férjhez megy!	
Ő nem tud eljutni.	elindulni
Öntesz nekem is könyvet?	kávét
Öntsünk tiszta vizet a köcsögbe.	pohárba
Öt pillanat vagy mi egy, egy pillanat...	
Pelikános, úgy értem flamingós.	
Racionálisan és ökonomikusan akarjuk megépíteni a világot magunk körül, hogy minél könnyebben szoronghasunk.	mozoghassunk
Rakjak borsófőzeléket, vagy babfőzeléket?	
Remélem, hogy nem találkozunk majd olyan rég.	soká
Rendeztem pár cuccot a H&M oldaláról.	rendeltem
Sokáig volt tudatlan, vagyis dehogy, eszméletlen.	

Száll egy kecske.	fecske
Szegény virág, már nem sokáig bírja. Pedig adtam neki virágot. Virágot? Vízet.	
Szerepelnek benne hímes híres színészek.	
Szétszakadt a vonal, vagy mi, megszakadt.	
Szia, Ildi, egy whiskast kérek!	whiskyt
Találkozzunk a Rákóczi körúton □ ja nem, a téren.	
Találtam egy penészes kenyeret □ nem! nem penészes! egész	
Tavaly vettem, izé tegnap.	
Történelmi pillanat volt a filmtörténelem- ben.	filmtörténetben
Ugyan már, a muslincák ártatlanok □ meg ártalmatlanok is!	
Van csuklód?	kulcsod
Van kés a fürdőben.	olló
Van olyan aranyér a hűtőben.	Piros Arany
Zombi, a drogkereskedő rendőrkutya □ drogkereső!	
Zsámbékról annyi zsömlét hozott.	szőlőt
Zsuzsa szokott zuhanyozni? Mit mondtam? Zuhanyozni? Dohányozni...	

## 2. Grammatikai hiba

### Megakadás

– Kinek vetted a szalonnát? – Lajosé □ nak.  
...ahol italokat szolgáltak ki.  
A jövő hetet pedig nem tudtam vele meg-  
beszélni.  
A Marson légtérhajlítással megyünk.  
A negyedik vizsgázik.  
A szövegkönyv 17. oldalán.  
Adok másik kanált  
Annyira nehéz voltak.  
Arról olvashatunk ezt a részt, amit nekünk  
olvasni kellett.  
Azt akartam felkérdezni vagyis megkér-  
dezni.  
Boldogok, mert nem vannak a szoba köze-  
pén □ nincsenek.

### Szándékolt közlés

szolgálnak fel  
hetet  
  
A Marsra  
negyedik  
oldalán  
kanalat  
annyira nehezek voltak  
ebben a részben



De ezt csak ez a gyártó gépei tudják?	ennek a gyártónak a gépei
Egy olyan témáról lesz most szó, ami eddig még soha.	amiről
Elviszed? Nem, nem belefér a táskámba.	nem fér bele
Érzed valami furcsát?	érzel
Ez egy új folyamat, amibe belül változik.	ami
Ezt erre is lehet ráalkalmazni, alkalmazni.	
Fel tudod olvasni párat?	fel tudsz olvasni
Fiatalokkal és idősebbel, idősebbekkel.	
Ja, ő utálja □ egymást Lacival.	ők utálják
Jobb, ha nyugaltságban van egész idő alatt.	nyugalomban
Két percre vagyunk otthonról, otthontól.	
Kicsit elkalandkodtam.	kicsit elkalandoztam
Le van törve a lámpái.	le vannak törve
Lement az aluljáróban.	aluljáróba
Megkérdezték, miben írom a diplomamunkámat.	miből
Menj a szobádból!	szobádba
Most nincsenek olyan sok ember.	
Múlt héten elmentünk két éttermekbe is.	étterembe
Nem akartam enni □ megenni előled.	
Nem annyira beazonosíthatóvá teszi...	nem teszi annyira beazonosíthatóvá
Nem fogok megenni.	fogom
Nem látom rajta, hogy akarna anyaként lenni.	anya akarna lenni.
Nem megy arra a vonat, mert javítják a sínet.	sínt
Nem tetszenek nekem ezek a sok betegségek.	Nem tetszik nekem ez a sok betegség.
Nem van, vagyis na □ nincs kedvem mosogatni.	
Nem véletlenül kerül oda, ahol.	oda, ahová
Nincs ehhez hasonló színek.	nincsenek
Nincsenek igazán sok belőle.	nincs igazán sok belőle
Óckodtam ettől a tartósabb műszempilláktól.	ezeiktől

Ott van a Dolomitokban nem messze □ aaa

Dolomitoktól nem messze.

Sokkal szépen hajtogatod.

tűztől pattant

Vagy a gépeink rossz, vagy az internet.

szebben

tűzről pattant

rosszak

### 3. Kontamináció

#### Megakadás

...egy helyre odacsoportosítva.

...és egy új verzióként születik újra □ újjá.

A belvárosban megint útleterelés van.

A hölgy megmondta, hogy ők fognak dönteni a kislány felett.

A kábítószer szerintem kifordította...

A kapcsolatomat be szeretném zárni.

A lakó fők száma...

A nyelv tudományozása a számítástech-  
nika szemszögéből.

A rendőrök a forgalmat kerelőútra terelték.

Az író a művet leveti, papírra veti.

Azért lásd be, mi két összeillő pár vagyunk.

Azt mondta, hogy tizenhét óráig ideér,  
vagy na, hét óráig.

celufólia

De a formé □ na, morféma csak a formára  
vonatkozik.

De ezt most mikroportba, őő mikrofonba  
vagy portba éneklede?

Emeld fel a keredet!

Én igyekszem minél kevesebbet foglal-  
kozni a politikával, mert különben bele-  
robbannék.

Én is szembefutottam már vele.

#### Szándékolt közlés

egy helyre csoportosítva x

egy helyre odatéve

újra x újjá

útlezárás x forgalomterelés

kislánnyról x kislány sorsa  
felett

megváltoztatta x kifordí-  
totta önmagából

le szeretném zárni x be sze-  
retném fejezni

a lakók száma x a lakásban  
lakó fők száma

nyelvtudomány tanulmá-  
nyozása

kerülőútra x terelőútra

leír x papírra vet

összeillő pár vagyunk x két  
összeillő személy/illető va-  
gyunk

tizenkilenc óráig x hét  
óráig

celofán x alufólia

morféma x forma

mikrofon x port

kezedet x karodat

felrobbannék x belerok-  
kannék

velem is szembejött már x  
én is belefutottam/ össze-  
futottam

Ez azonban nem róható fel a fordító számlájára.	nem írható fel a számlájára x nem róható fel a fordító- nak
Ezzel a gyerekekkel csak a gaj van!	gond x baj
geopárd	gepárd x leopárd
Harmincnyolc pontot lett konkrétan.	pontot szereztél x pont lett
Hívni kell a vízvezetőt.	vízvezeték x szerelőt
Hol az ágypléd?	ágytakaró x pléd
Hülyeség, amit mondolsz.	gondolsz x mondasz
Ismerek egy olaszt, 13 éve él itt Pesten, tud magyarul és van egy pizzériázója.	pizzázó x pizzéria
Lejössz velem kennyi?	kenyeret x venni
Mit dudál ez a tapó?	mit dudál ez a tapló/tahó
Na mi is egy cipőben evezünk!	egy cipőben járunk x egy csónakban evezünk
Nagyon zokon esett, amit mondott rólam.	zokon vettem x rosszul esett
Nem tudott itt jönni.	nem tudott jönni x nem tud itt lenni
Olyan butálékos.	tudálékos x buta

#### 4. Téves kezdés

##### Megakadás

...ugyanis ért-félreérthető volt...

...ugye el is kez □ pontosabban folytatódott...

A 3-as kanyarban majdnem eldobta a seg □ hátulját az autó.

A bal egy agyféltekében

A Csabi szí hívott.

A hibákat úgy ve, úgy javítják...

Alig vártam, hogy megszadul, megszabaduljak tőle.

Ameddig nem történik meg, addig nem fo, addig nem tudunk lépni.

Egy ilyen eszte □ aszteroidát kérnék!

El fogom költ □ küldeni neked a feladatot.

És ők nem növelhetik az áru, az árat?

Este lef □ elválás előtt pedig moss mindig fogat.

Ettől a hidegtől bőröm kiszárad és ilyen ilyen seee sérülékennyé válik

Hol vagytok, a bá há backstage-ben?

Jó volt rá a húsz százalékos ak kedvezmény.

Mégiscsak megismerhet ööö megismételhető.

Most megsérül □ megsértődöttél rám?

Nem, kata.. baszkok írták.

Nyisd ki a szá a szemed.  
 Persze, tudjuk no □ foglalkoztatni.  
 Rakd a helyére a szege □ székeket kérlek!  
 Rögtön az jutott eszembe, hogy Kra, Kaliforniában, tudod, hogy mennyi fok van.  
 Szeretem a különleges ízű teákat, a hibikszu, a hibiszkusz a kedvencem.  
 Utána lehet kés □ kérdezni

### 5. „Nyelvem hegyén van” jelenség

Megakadás	Szándékolt közlés felelős
A kérdésben ööö nem ööö nem én vagyok a felettes ööö felelés.	
A Tündérfürt utcai gyógyszergyárban voltam, na, tudod □ nem dolgozni, hanem mikor tanulsz is □ izé gyakorolsz □ gyakorlaton!	
Add ide az □ na, távorientált. Izé, irányított.	távirányító
Adj egy izét □ csókot.	
Akik ilyen metrószok, kukások □ na azok a □ tudod...	közterület fentartók
Amikor megkaptam a diplomát □ nem az érettségit □ na, a bizonyítványt.	
Anti □ ante aneszteziológus	
Aztán vettem egy olyan izét is, mi az □ na, amivel a téstát, na az a ...	szűrő
Bekapta aaa a izé a legyet.	
Blakett, plakett, na plakát már basszus!	
Én úgy vettem észre, hogy □ amikor □ a □ hogy is hívják?	lift
És akkor felveszed azt az izét □ Tudod, ami olyan, mint egy hajpánt, t-vel kezdődik □ Na, mondd már!	tiara
Ez ilyen ilyen vizes fék?	hidraulik
Ez most nem a, tudod, az a zsidó szokás, ami □ jom kipur, igen.	
Hol verjük fel aaa kempinget, mármint na, a sátrakat?	
Ilyen izé van benne, töltölék izé, töltelék, na.	

Jaj, ott van egy □ egy □ mindegy, már elment.	pillangó
Kimentem oda □ tudod □ kint □ a zöldségek.	kertbe
Már tegnap átküldtem a macskáról a videót, hangizét.	hangfelvételt
Nálad van a □ na, tudod, amire a számokat is rá kell írni, ilyen kis igazolványos izé □ Havonta veszed, tudod te mi az, csak már fáradt vagyok, nem jut eszembe.	bérlet
Nehézkes lenne elhozni magam, magamban, na hogy autó nélkül.	egyedül
Nem adta ide a falk □ flakon, tudod □ üveget.	
Nem tudtok mentőt hívni? Nálam nincs, na mi az a kezedbe, rádió □ rádiótelefon azokkal a hullámokkal.	mobiltelefon
Olyan gyors volt, mint a □ a □ tudjátok a leggyorsabb állat □ az a pöttyös □ áh hagyjuk!	gepárd
Ő is benne lesz az osztálytáncban, mert valami madár lesz, pingvin vagy pávián □ Ja tudom, flamingó!	
Össze fo fo fo fognak vagy mennek a szálak.	összefonódnak, egymásba fonódnak
Sokszor egy ilyen, segítsetek! Mi az a □ közbevetés! Közbevetésnek hívják.	
Úgy tűnsz, mint aki □ izé □ szomorú.	
Ugyanaz a plagi plagi plagi na mi az plagizálás.	
Úristen, de degradált □ nem □ deg □ hogy is mondják?	degenerált
Valahogy odaküldöm a hogy mondják szellememet.	
Vár a szervpranszlatáció □ szervtraszplatáció □ szervátültetésre, na!	
Vedd le légydzi a bödönt □ vagy mi az □ lábost.	

**6. Perszeveráció****Megakadás**

A békülésben szerepe volt Fanniban □  
nak.  
A múltkor is az almát szénné égette, mint  
az álmat.  
A nyelv olyan viszonyban van a lélel,  
mint a felhő az éggel.  
A WWF az egy önkéntes tervezet?  
Az az olasz is annyira jól tud olaszul □  
vagy na, angolul.  
Az EEG- gép az egy agyi gépalkotó mód-  
szer.  
Az embernek rövidebb ideje van  
monitoriz monitorozni a beszédét.  
Belement a szemembe a szem.  
De nem tudom, hogy kapok majd szabít  
nyárra, mert janyuárban mondok fel.  
Egyet fizet, kettőt fizet.  
El fogsz elesni, ha nem vigyázol!  
Elmostam a kávévőzőt.  
Én is a Microsoft szövegszörkesztőt □  
szövegszerkesztőt szoktam használni.  
És kivel veszéltél?  
Este van, ne menj már a voltba.  
Eszembe sze jutott illet mondani.  
Fel van szelelve, van hűtő, mosógép.  
Felvázolom a vázlatot, vagyis a helyzetet.  
Használtam a fürdősőt.  
Két tojást törtem fel, hogy aztán azt törje  
bele a serpenyőbe.  
Kétszer is kétszerített.  
Kicsit túl tettem magam.  
Milyen fa volt? Nem tudom, hárs vagy  
nyárs.  
Most ilyen viktoriánus-kori brit lord hajod  
van.  
Muszáj áldozdatot hoznom.  
Ott talizunk Talízzal.  
Önök imádnak idézőjeles szöveget idézni.  
Szedjete még a levelesből.

**Szándékolt közlés**

Fanninak  
állat  
léttel  
szervezet  
képalkotó  
  
por  
januárban  
vihet  
el fogsz esni  
kávéfőzőt  
  
beszéltél  
botba  
eszembe se  
szerelve  
fürdősót  
tegye bele  
kétszer is kényszerített  
kicsit túl ettem magam  
nyár  
hajad  
áldozatot  
Alízzal  
írni  
levesből

újfujlandi  
Vak tyúk is tyalál szemet.

újfundlandi  
talál

## 7. Anticipáció

### Megakadás

...ekkor hirtelenül megjön a férj hirtelen.  
A világügyi egészség □ vagyis Egészség-  
ügyi Világszervezet most kiadott jelen-  
tései szerint...  
Az egy németkutyász kutya.  
érleléselemélet  
Ez ilyen régi rég délelgetett álmom, hogy  
legyen egy ilyen pici lakásom, és ara-  
nyos.  
Ilyen frodros ruhában jövök.  
Ilyenkor hiányzik az alany felismerését  
szegítő szerkezet.  
Kérek egy csült csikeszárnyat.  
Két felsőt fizettem, és csak kétezret fizet-  
tem.  
Kívánom, hogy lobb legyél!  
Lányatok mondjátok már meg, hogy...  
Meggérted téged.  
Mekkenj már meg!  
Mi vaj van a vajjal?  
Nagyon baitta magát Balázs.  
Nagyon szok szöveget kell elolvasnom.  
Nem bárom bár...  
Nem megyünk halba! Vagyis étterembe.  
Halat sütök este.  
Pendelünk pizzát?  
Sok a munkaerő most a programbunkban.  
tükerszimmetrikus  
Ugyanannyira tél □ tér el...

### Szándékolt közlés váratlanul

Az egy németjuhász kutya.  
érveléselemélet  
dédelgetett

fodros  
segítő szerkezet

kérek egy sült  
csirkeszárnyat  
két felsőt vettem

jobb  
lányok  
megkértelek  
Mukkanj  
baj  
leitta magát Balázs  
sok  
nem bánom, bár

rendelünk  
programunkban  
tükörszimmetrikus

**8. Metatézis****Megakadás**

...cseret helyélhettek...

A nyelvem van hegyén jelenség.

Amit ilyenkor lenni tehet tenni lehet.

Attila miért ennyire gát a legényen?

Az autópályán kamarbol volt.

Azt hiszem, a kognitív elméleten apalul.

Betettem a szekrényt a kabátba.

Bocsánat, nem tudom apát adni a telefonhoz, épp lapat havátol.

Csak aztán nem bírta a gyomrát a politika, vagy a politikát a gyomra.

Csekken voltam postát fizetni.

Elbúcsúztatjuk az újévet és köszöntjük az óévet...

Elhagyja a nyakát a fejről.

Elmennél vegyet jenni, na □ jegyet venni, kérlek?

Éppen az andorrai vereség napján nyújtották be a költségvetésnek □ országgyűlésnek a költségvetésre vonatkozó javaslatot.

Fosdd meg a mogadat.

Hé, nem láttad az üvegszemem?

Hegyél tejre ööö tegyél helyre.

Jó ez a szőrfesztivál □ vagyis sörfesztivál...

két kimeneteles az esélye

Köszönöm, megkaptam a válaszomra a kérdést.

Láttam én már varón karjút.

Már megint lekrám van!

Mi a szócs?

Nekiment a fával a biciklinek, jajj dehogy, a biciklivel a fának.

Nem akarok reszet sajtani, vagyis sajtot reszelni.

Nem akarsz feletonálni?

**Szándékolt közlés**

helyet cserélhettek

Attila miért ennyire legény a gáton?

karambol

alapul

kabátot a szekrénybe

havat lapátol

Postán voltam csekket fizetni.

fordítva

elhagyja a fejét a nyakáról

Mosd meg a fogadat. szemüvegem

kétesélyes a kimenetele a kérdésemre a választ

karón varjút  
reklám  
csősz

telefonálni



Nem haragszom, csak nem dutom ☐ tudom  
megérteni.

Nem porbléma.

Ninden map leviszem a kutyát sétálni.

Szeretnék egy borítékos bélyeget.

tespi

Zárd be a fürdőt a macskába!

probléma

minden nap

bélyeges borítékot

tespi

macskát a fürdőbe

### 9. Egyszerű nyelvbtlás

#### Megakadás

...tevénap az történt...

A harmincegyes út felújíjtása vagy  
felújítása.

A mű érdeme, hogy az elvont témát  
könnyen megragadhatóvá tette.

A szőkincs, akarom mondani a szó-  
kincs fejlesztéséhez...

Amikor megszületik egy gyarak gye-  
rek...

Az még egyáltalán nem feltételezi hogy  
☐ feltételezi...

Bepurnyadtam ☐ vagyis  
bepunnyadtam.

Egyesül vagyok.

Grízishelyzet ☐ krízishelyzet esetén...

Halat süttem ööö sütöttem.

Imádom a lókákat.

Jó napot kívánok! Jegeket, bérleteket  
kérem szíveskedjenek felmutatni!

Kérsz hozzá ubokát is?

Kortás irodalom.

megvalószul

Miért nem tartanél?

Mikor restaurálták?

Minyimálpárok

Mit kell csinálnom, könyörögni,  
pitini?

Nagyon dolgos vagyok mostanság.

Nem tudom mi vajod.

Ő is ezt mondta, miután  
kitelepetítették.

#### Szándékolt közlés

megragadhatóvá

egyedül

rókákat

jegyeket

uborkát

kortárs

megvalószul

tartanél

restaurálták

Minimálpárok

pitizni

dolgos

bajod

kitelepítették

Pusztán arról van csak szó, hogy nem  
veszel komolyan. van csak szó

Teljes egy óra volt, mire elkészült a  
sempillám, na szempillám.

#### 10. Több típusba sorolható jelenségek

##### Megakadás

##### Szándékolt közlés

A költözéshez kellenek neked koboznod, izé, do-  
bozok?

Állatoz ☐ nem, nem állatoz ☐ utánozd az állatoz  
hangját!

Az húlt ho nyúlt hosszabbra.

Kily milyen csaj hívott most?

Most megeszem ezt a kólás hülyeségeket.

Ne ezt, ez olyan vanyadt!

nyamvadt

Szerintem most ebé, ebédelini? Reggelizni fogok.

## KÖNYVISMERTETŐK

**Bóna Judit:**

*Női beszéd – férfi beszéd a fonetikai és pszicholingvisztikai  
vizsgálatok tükrében*

**Akadémiai Kiadó, Budapest, 2016.**

A nők és a férfiak beszédbeli, nyelvhasználatbeli különbségeivel több tudományterület is foglalkozik: szociolingvisztika, pszicholingvisztika és fonetika. A kötet bevezetője átfogó képet ad ezen kutatások főbb eredményeiről. A szociolingvisztikai kutatások kimutatták, hogy a nők nyelvhasználata általában választékosabb, udvariasabb és konzervatívabb; ellenben könnyebben képesek alkalmazkodni az újabb nyelvi változatokhoz. A nők és a férfiak társalgási szokásai is eltérőek: a nők inkább kérdeznek, mint kijelentenek; partnerorientált és kooperatív attitűd jellemzi őket. A pszicholingvisztikában is számos szempontból vizsgálható a nők és férfiak beszédének eltérése: beszédtervezési sajátosságok, olvasás, idegennyelv-tanulás, anyanyelv-elsajátítás és nyelvi zavarok előfordulása. A beszéd akusztikai vizsgálata szintén rámutat a női és férfi beszéd különbözőségére: az eltérő alaphangmagasság laikusok számára is jól érzékelhető, különbségek lehetnek azonban a zöngeminőségben, intonációban, tempóban stb. A két nem közti eltérések egyrészt biológiai okokra vezethetők vissza, de a társadalmi különbségek is jelentős befolyással bírnak a beszédre.

A kötet a női és férfi spontán beszéd hasonlóságait és különbségeit mutatja be objektív akusztikai fonetikai mérések és pszicholingvisztikai elemzések segítségével. Három életkori csoportot vizsgál: gyermekeket (9–10 évesek), fiatal felnőtteket (20–32 évesek) és időseket (66–90 évesek). Az egyes életkori csoportok összevetése lehetővé teszi, a nemek közti különbségeken túl az életkori eltérések bemutatását is. Összesen több mint 40 000 adatot dolgoz fel különböző statisztikai módszerekkel. Elemzi a beszédprodukciónak a stratégiákat, a beszéd temporális kereteit, illetve a beszéd érthetősége szempontjából legfontosabb beszédhangok (magánhangzók és felpattanó zárhangok) akusztikai-fonetikai sajátosságait. Szándékoltan olyan kérdéseket vizsgál, amelyek korábban kevésbé jelentek meg a női és a férfi beszéd összevető vizsgálatában, és nem célja olyan sokat kutatott és sokak által bemutatott kérdések elemzése, mint például az alaphangmagasságbeli különbségek.

A beszédprodukciónak a folyamatok leginkább a megakadásokon keresztül elemezhetők, ezért a különböző típusú megakadásjelenségeket (bizonytalan- és hiba típusú megakadások) vizsgálja gyakoriság, illetve fonetikai saját-

tosságok (például időzítés) szempontjából. Általánosságban elmondható, hogy az életkornak jelentősebb hatása volt az eredményekre, mint a nemnek. A nem csak néhány esetben befolyásolta szignifikánsan a vizsgált paramétereket. A nemek között szignifikáns különbség volt kimutatható a megakadások percenkénti előfordulásában az idősek körében: az idősek nők több megakadást produkáltak adott idő alatt. Csak a bizonytalansági megakadásokat nézve (hezitálások, nyújtások, ismétlések, újraindítások, szünet a szóban és töltelékszók), szintén szignifikáns különbség volt kimutatható az idősek nők és férfiak között a gyakoriság (percenkénti előfordulás) tekintetében. A gyerekek csoportján belül a szóban tartott szünetek időtartamát befolyásolta szignifikánsan a nem: a fiúk szignifikánsan hosszabb szüneteket tartottak. Az életkor hatása több esetben eltérő módon jelentkezett nőknél és férfiaknál.

A megakadások nem függetlenek a beszéd temporális jellemzőitől. A beszéd temporális sajátosságai közé sorolható a beszédtempó, az artikulációs tempó és a szünettartás; ezek összefüggésben állnak a tematikus egységek, társalgási fordulók és beszédszakaszok időtartamával is. A női és férfi beszéd kapcsán több sztereotípiát is ismert: a nők többet és gyorsabban beszélnek, mint a férfiak. Ezen közvélekedéseket nem feltétlenül igazolják tudományos kutatások, több esetben eltérő eredményekről olvashatunk. A kötet szakirodalmi áttekintője rávilágít ezekre a különbségekre, és tovább keresi a választ a felmerült kérdésekre a következő paraméterek mentén: társalgási fordulók hossza, tempóértékek (beszéd- és artikulációs tempó), beszédszakaszok hossza, tagoltság, illetve szünettartás (szünetek aránya, időtartama és szünettípusok: néma, kitöltött és kombinált). Az eredmények szerint a társalgási fordulók időtartamát és szószámát sokkal inkább az életkor befolyásolta, mint a nem: a gyermekek társalgási fordulóit a legrövidebbek, az időseké a leghosszabbak. A tempóértékek tekintetében szintén az életkor befolyásoló hatása a jelentősebb, a nemek között csak a fiatal felnőttek csoportjában volt szignifikáns különbség: a férfiak artikulációs tempója nagyobb volt. Nem igazolódtott tehát az a közvélekedés, miszerint a nők gyorsabban beszélnek. A beszédszakaszok hosszát összevetve, csak a fiatal felnőttek eredményei között volt szignifikáns különbség: a nők beszédszakaszai átlagosan hosszabbak voltak. A szünetek időtartamát tekintve szignifikáns különbség volt mérhető a fiatal felnőttek és az idősek csoportjában is a két nem között: a férfiak beszédében átlagosan a hosszabb szünetek voltak jellemzők.

A kötet szegmentális szinten is vizsgálja a női és férfi beszéd különbségeit. Külföldi szakirodalmak szerint a nők és a férfiak beszéde között elsősorban a magánhangzók és az explozívák akusztikai szerkezetében vannak különbségek, ezért a jelen kötet is ezen beszédhangok akusztikai szerkezetét elemzi. A magánhangzók esetében elsősorban az időtartamot és a formánsszerkezetet szokás vizsgálni. Mivel a szakirodalomban gyermekkorról több adat is van már magyarra vonatkozólag, a jelen kötet a fiatal felnőttek és az idősek csoportjában vizsgálja a nemek közti ejtéskülönbségeit. A korábbi szakirodal-

makkal egyezően az idők hosszabb magánhangzó-időtartamokat produkáltak, és a nők formánsai magasabb frekvenciákon realizálódtak. A fiatal férfiak fonetikailag is párt alkotó rövid-hosszú magánhangzói között kisebb különbség volt mérhető, és a magánhangzók átlagosan kisebb magánhangzótérben realizálódtak, mint a több csoportban. Az explozívák esetében a zöngekezdési időt (VOT) szokás elemezni. A kötet a fentebb említett három életkori csoportban vizsgálja a zöngétlen zárhangok (*p*, *t*, *k*) zöngekezdési idejének alakulását a nem függvényében. A zöngekezdési időt több tényező is befolyásolja: az artikuláció helye, a követő magánhangzó minősége, a beszédtempó, a beszédtempó stb. Korábbi szakirodalmak szerint a férfiak zöngekezdési ideje rövidebb (viszont az artikulációs tempójuk is gyorsabb). A kötet eredményei egyeznek az eddigi szakirodalomban leírtakkal, a gyermekek esetében azonban még nem alakult ki egyértelműen különbség a lányok és a fiúk ejtésében. A mássalhangzó képzési helye mindegyik életkori csoportban jelentősen befolyásolta a zöngekezdési időket.

A kötet további kutatási irányokat is megemlít, érdemesnek tartja az elemzett beszédanyagok szövegnyelvészeti vizsgálatát és tartalomelemzését annak kiderítésére, hogy vajon a hosszabb összefüggő beszéd milyen információsűrűséggel jár együtt, van-e különbség a nők és a férfiak, illetve az eltérő életkorú beszélők beszéde között.

A kötet eredményei több gyakorlati területen is felhasználhatók. A nyelvészeti alkalmazásokon túl (szociolingvisztikai, pszicholingvisztikai, fonetikai vonatkozások és kommunikációtan) alkalmazhatóak a kriminalisztikában és a beszédtechnológiában is.

*Huszár Anna*

**Szántó Anna:**  
***Beszédfeldolgozási folyamatok változásai a fejlesztés függvényében***

**Takács Etel Pedagógiai Alapítvány, Budapest, 2017.**

Az olvasás, írás elsajátításának fontos előfeltétele az életkornak megfelelő anyanyelvi képességek megléte. Ezalatt nemcsak a gyermek érthető, tiszta és grammatikailag helyes beszédének kialakulását értjük, hanem a szavak, mondatok, szövegek pontos feldolgozását is. Ide tartozik a beszédészlelés és a beszédmegértés összetett folyamata, amelynek jelentőségét nem lehet eléggé hangsúlyozni. A beszédpercepció elmaradások negatívan befolyásolhatják az írott nyelv elsajátítását, a szövegértést és az idegen nyelv tanulását is, ezért rendkívül fontos időben feltárni és fejleszteni az elmaradást mutató területeket. Szántó Anna disszertációjában óvodások és első osztályos gyermekek beszédpercepció képességeit vizsgálta, majd a vizsgálati csoportba tartozó

óvodások és iskolások, hét hónapon át tartó fejlesztésen vettek részt, Gósy Mária – Imre Angéla (2007): *Beszédpercepció fejlesztő modulok* című könyve alapján, a szerző által összeállított fejlesztési tervekből.

Az előszó utáni bevezetésben részletes összefoglalót olvashatunk a beszédészlelés és beszédmegértés működésének folyamatáról, fejlődéséről és annak hibás működéséről. A fejezet végén a szerző bemutatja azokat az eljárásokat, teszteket, amelyek a hazai gyakorlatban is használatosak a beszédpercepció különböző részterületeinek vizsgálatára.

A kutatás egyrészt kontrasztív, mert különböző korosztályok egy időben nyert vizsgálati eredményeit hasonlítja össze, másrészt longitudinális, mivel ugyanazon gyermekek beszédpercepció képességeit vizsgálja újra egy bizonyos idő elteltével. A kutatásban nagycsoportos óvoda és első osztályos gyermekek beszédészlelési és beszédértési teljesítményét vizsgálta a szerző a GMP-diagnosztikai eljárás 14 résztesztjével. A GMP beszédpercepció diagnosztikai eljárást dr. Gósy Mária fejlesztette ki, amelynek célja 3–14 éves gyermekek beszédpercepció elmaradásának, zavarának és ezek mértékének a feltárása. A vizsgálatban összesen 56 adatközlő vett részt, 28 óvoda és 28 iskoláskorú gyermek, akik azonos területen lévő két óvodából és egy iskolából kerültek ki. A kutatás első felére az őszi tanévben került sor, majd mindkét korosztályi csoportból 14-14 gyermek beszédpercepció fejlesztésben részesült hét hónapon át. A többi gyermek nem vett részt foglalkozáson, ők a kontroll csoportba tartoztak. A kutatás második szakasza a tavaszi félévben történt meg, ahol újra felmérték a gyerekeket.

Az év eleji felmérések azt az eredményt hozták, hogy minden korcsoportban a fonetikai észlelés és a beszédhangok azonosítása működött a legjobban a tesztmondatok ismétlésében. Mindkét korcsoportban a fonológiai észlelés átlagos teljesítménye jóval alatta maradt a fonetikai észlelésnek, azonban az életkor előrehaladtával (iskolásoknál) a két folyamat jó működése közötti különbség csökkent, ami azt jelentette, hogy a fonológiai észlelés egyre jobban megközelítette a fonetikait. Mindezek mellett a fonológiai észlelés átlagos teljesítményei mind az óvoda, mind az iskola korosztályban három év elmaradást mutattak.

A vizuális észlelésben és a ritmusészlelésben az életkorban elvárthoz közeli átlagteljesítmények születtek. A szeriális észlelésben mindkét korcsoportban több évnnyi elmaradás volt tapasztalható, és a beszédhang differenciálása még az iskolásoknál is nehézséget okozott.

A szövegértés mindkét csoportnál alacsony szinten működött, a mondatmegértési feladatot a gyermekek sikerebben oldották meg, főként az iskola korosztályban. A vizuális és verbális memóriát vizsgáló tesztben a két korcsoport gyermekei az elvárt eredményeket hozták. A szóaktiválás feladatban az aktivált szavak számában nem találtak különbséget a két korosztály között, ami az óvodásoknál az elvárthoz közeli eredmény, azonban az iskola korosztályban ez az érték egy évnnyi elmaradást mutatott.

Az első teszt felvételét követően került sor – a *Beszédpercepció fejlesztő modulok* című könyv alapján speciálisan nagyobb létszámra összeállított – fejlesztési tervek megvalósítására. Összesen 27 hétre készültek el a tervek, heti bontásban, amely a kötetben a 218. és a 294. oldal között található. Az óvodában a nap folyamán bármikor fejleszthették a gyermekeket, az iskolában napi 5 perc gyakorlási időt tudtak szánni a fejlesztésre. A foglalkozásokat a teljes óvodai, illetve iskolai csoportban tartották meg az óvónők, illetve az iskolai tanítók.

A fejlesztést követően újra elvégezték a GMP-diagnosztika tesztjeit a két korcsoportban. Az óvodásoknál a vizsgálati és a kontroll csoportban is statisztikailag igazolható volt a javulás az akusztikai-fonetikai és a fonológiai észlelésben az év eleji eredményekhez képest. A szeriális észlelés tekintetében csak a vizsgálati csoport esetében igazolható a fejlődés, azonban a gyermekek ennek ellenére sem érték el az életkornak megfelelő eredményt. A szótalálási tesztben is kimutatható volt a vizsgálati csoport fejlődése.

Az iskolás korosztályban szignifikáns javulás volt tapasztalható a vizuális észlelés és a szeriális észlelés feladatokban. A beszédhang-differenciálás próbában mindkét csoport eredményei fejlődést mutattak, azonban a többi területen nem volt statisztikailag kimutatható különbség a két csoport eredményei között.

Összehasonlítva a fejlesztésben részesült óvodások és fejlesztésben nem részesült iskolások eredményeit arra jutott a szerző, hogy szignifikáns különbség volt a szövegértés tekintetében a két csoport között. A fejlesztett óvodások jobb eredményt értek el a nem fejlesztett iskolások csoportjához képest. Az eredmény azért rendkívüli, mert a szövegértés a teljes beszédmegértés folyamatának működését vizsgálja, azaz minden észlelési szintet magába foglal.

Az eredmények tükrében kimondható, hogy az iskolai tanulmányok megkezdése előtt egy évvel fontos volna a beszédfeldolgozási szintek működéséről képet kapnunk, majd az elmaradott területeket céltudatosan fejleszteni. Ebben nagy segítséget nyújt e kötet, melyben kidolgozott fejlesztési terveket kapunk, amely jól használható nagyobb csoportok fejlesztése során.

A kiadványt nem csak nyelvészeknek ajánlom, segítséget nyújthat óvónőknek, tanítóknak, gyógypedagógusoknak, illetve logopédusoknak is – minden olyan szakembernek, aki az iskolakezdési időszak előtt fejleszti, segíti a gyermekeket képességeik kifejlődésében.

*Désfalvi Ildkó*

**Bóna Judit (szerk.):**  
***Új utak a gyermeknyelvi kutatásokban***

**ELTE Eötvös Kiadó, Budapest, 2017.**

Az anyanyelv-elsajátítás vizsgálata a 20. században került a nyelvtudomány művelői érdeklődésének a középpontjába – írja a kötet Előszavában Bóna Judit, az *Új utak a gyermeknyelvi kutatásokban* szerkesztője. Azóta a gyermeknyelv kutatása számos változáson ment keresztül, a gyermeknyelvi naplótól eljutottunk a nagy méretű beszédatadabázisokig, valamint olyan új kísérleti módszerek alkalmazásáig, mint az elektromágneses artikulográfia vagy az ultrahang. A könyv írásain keresztül nyomon követhetjük ezt a fejlődési folyamatot, valamint kijelölhetjük magunk számára a lehetséges további kutatási irányokat.

Az Előszó után Gósy Mária összefoglalását olvashatjuk az anyanyelv-elsajátítás és a gyermeknyelv kutatásának történetéről, elméleteiről, módszereiről. A huszadik század elején a kutatókat elsősorban a szókincs kérdésköre érdekelte, ezt követte a hangtan, a morfológia és a fonológia feltérképezése, majd egyre inkább megjelentek nemcsak a produkcióra, hanem a percepcióra irányuló eredmények is. A tipikus nyelvfejlődés mellett egyre több kísérlet irányul napjainkban is az atipikus fejlődési mintázatokra, a két- és többnyelvűség, a jelnyelvhasználó közösségek nyelvhasználatának, valamint a pragmatikai tudás szerepének leírására is. A hangrögzítés elterjedésével lehetővé vált a szupraszegmentumok elemzése, valamint longitudinális, illetve különböző akusztikai-fonetikai kutatások elvégzése is. Napjainkban a beszédatadabázisok korát éljük, a magyar nyelvű gyermeknyelvi adatbázisok közül a legjelentősebbek a GABI, a TiniBEA, valamint a MONYEEK.

A könyv második fejezetében Bóna Judit hívja fel a figyelmet a beszédatadabázisok, valamint a GABI (Gyermeknyelvi beszédatadabázis és Információtár) jelentőségére. A legismertebb nemzetközi adatbázisok a TIMIT, a Switchboard, a Verbmobil vagy a Kiel Corpus of Spontaneous Speech. A nemzetközileg is jelentősnek számító BEA (BEszéltnyelvi Adatbázis) magyar köznyelvi hanganyagokat tartalmaz. A gyermeknyelvi beszédatadabázisok között a legismertebb a CHILDES (Child Language Data Exchange System), amely 26 különböző nyelvből tartalmaz felvételeket. A GABI fejlesztése 2013-ban kezdődött az ELTE Fonetikai Tanszékén. Két nagy osztályba sorolja a hangfelvételeket, a 3–9, illetve a 9–18 évesek életkori csoportjába, meg felvételek külön protokoll szerint készülnek. Jelenleg 405 db hanganyagot tartalmaz a korpusz, a cél 900 felvétel elkészítése. Elsősorban egynyelvű, tipikus fejlődésű gyermekek hanganyagait rögzítik, de található néhány atipikus, illetve kétnyelvű gyermek felvétele is az adatbázisban. Előnye, hogy nagy számú adatközlővel többféle beszéd típusban tartalmaz felvételeket, így az univerzális beszédjellemzők kutatására is kiválóan alkalmas.



A következő fejezetben Vakula Tímea és Váradi Viola látja el az olvasót hasznos tanácsokkal a GABI-felvételek lejegyzésének tapasztalatai alapján a felvételkedés és az annotálás tudnivalóiról. A szerzők kiemelik, hogy gyermekek esetében több nehézség merül fel a felvételkedés során, mint a felnőtteknél. Természetes közegekben (iskola, óvoda, otthon) vesznek fel a hanganyagokat, hogy minél kevésbé legyen számukra szokatlan a szituáció. A spontán narratíva során dicsérettel, mosolygással szükséges biztatni a fiatalabb gyermekeket, növelve ezzel beszédkezdőket. Hasznos módszer emellett a kiegészítendő kérdés, a választós kérdés vagy az utolsó frázis megismétlése is. Nehézséget jelenthet fiatalabb életkorokban a pontatlan artikuláció és a különböző beszédhibák is, mindezek miatt az adatbázis fejlesztői tréningeket is tartanak a lejegyzők és a felvételkedítők számára egyaránt.

Markó Alexandra és munkatársai írásában a gyermeknyelvi artikuláció vizsgálatának új, műszeres vizsgálati módszereiről olvashatunk. Ezek az ultrahang, az elektromágneses artikulográfia, valamint az elektroglossográfia. Előnyük, hogy az artikulátorokat mozgás közben is nyomon tudják követni, valamint használatuk minimális kellemetlenséggel jár az adatközlő számára. Az ultrahang a nyelv felszíni kontúrját teszi láthatóvá. Az elektromágneses artikulográfia (EMA) módszere során a képzőszervek egyes pontjaira szenzorokat helyeznek, ezek érzékelik azok helyzetét és elmozdulását. Az elektroglossográf (EGG) a zöngeminőség vizsgálatára lett kifejlesztve, elektrodáinak segítségével a hangszalagok összetalálkozása és összeérése becsülhető meg. A gyermeki artikuláció vonatkozásában fontos kutatási irány lehet a jövőben a koartikuláció vizsgálata ultrahanggal és EMA-val, a mássalhangzók képzéshelyének megvalósulása vagy az ultrahangnak ajakvideóval való kiegészítése is.

A következőekben Horváth Viktória kutatásának eredményei olvashatóak a 6, 7, 8, illetve 9 éves gyermekek spontán beszédében előforduló megakadásjelenségek típusairól és időzítési sajátosságairól. Minden életkorban a kitöltött szünetek száma volt a legmagasabb, ezt követte a téves kezdések, valamint az ismétlések előfordulása. Az ismétlések a 6 és 7 éveseknél kivétel nélkül funkciószavakat érintettek, az idősebbeknél ez az arány kisebb. Az életkor előrehaladtával csökken a szó belsejében tartott néma szünetek aránya, valamint 9 éves korra a grammatikai hibák száma is. Az ismétlések szerkesztési szakaszai voltak a leghosszabbak minden korcsoportban. A csoportok között nem volt adatolható szignifikáns különbség a megakadások előfordulásában és temporális sajátosságaiban (a kitöltött szünet kivételével).

A szókincs gyarapodásának, minőségi változásainak kérdéskörével foglalkozik Neuberger Tilda. A gyermeki nyelv elsajátításának szakaszai a gögicselés, a gagyogás, a holofrázisok, a kétszavas közlések, a telegrafikus beszéd, majd hároméves kor környékén megjelennek az összetett mondatok, valamint három-négyéves korban az ún. verbális mámor időszaka figyelhető meg, amikor az igék száma hirtelen és nagymértékben megszorodik. 3–6

éves kor között jellemzőek továbbá a túláltalánosítások, a népetimológia, a metaforák, megszemélyesítések vagy a kontaminációk, amelyeket ebben az életszakaszban a gyermekek nem stilisztikai célból használnak. Szófajok tekintetében, 6–11 éves kor között a főnevek dominálnak a beszédben, 13 éveseknél azonban már az igék. Az ilyen vizsgálatok jelentősége, hogy segítségükkel fel lehet tární a megfelelő életkori szinttől való elmaradást, majd megtalálni az adekvát módszert a fejlesztésre.

Krepsz Valéria és Gósy Mária az anyanyelv-elsajátítás egyes életkori szakaszainak morfémaidőzítési sajátosságai kapcsán végzett kutatásaikról számolnak be. Kísérleti eredmények szerint az egyes toldaléktípusok a nyelvfejlődés eltérő szakaszaiban jelennek meg. A gyermeknyelvben a szótó/toldalék arányok különböznek a felnőttnyelvben tapasztaltaktól. A kiegyenlítődési tendencia értelmében minél több szótagból áll egy szó, annál inkább csökken a szótagok időtartama. A jelen kísérletben azt találták, hogy ez a tendencia a két és három szótagból álló tövek esetében a 4, az 5, illetve a 6 éveseknél is bekövetkezik, az annál hosszabb szavak esetében azonban csak a hatéveseknél. A toldalékok időtartama a négyéveseknél volt a leghosszabb, a hatéveseknél a legrövidebb, ami összefüggésben áll a kognitív funkciókkal és az artikulációs tempóval. További eredmény, hogy az anyanyelv-elsajátításban rendszerint korábban megjelenő toldalékok (pl. *-ban/-ben*) kiejtési időtartama rövidebb, mint a később megjelenőké.

Tar Éva beszédhanghibás tüneteket mutató gyermekek beszédében vizsgálta a szó eleji elől képzett zárhangok zöngésségi jellemzőit. Megállapította, hogy az elől képzett zöngés zárhangok előzöngés megvalósulásainak előfordulási aránya a tipikus fejlődésű és a beszédhanghibával rendelkező csoportban (BHH) szignifikánsan nem különbözött. A BHH csoportban jelentős egyéni különbségek voltak az előzöngés megvalósulások gyakoriságában. A zöngés-zöngételn kontraszt elsajátításának folyamatában a BHH-sok egy részénél a folyamat megrekedt abban a szakaszban, amikor még nem alakult ki a zöngésségi kontraszt. Akiknél viszont elsajátított a kontraszt, azok stabilabban használják a beszédükben, mint az azonos fonológiai fejlettségű, de fiatalabb gyermekek.

Ausmann Anita 7, 9, 11 és 13 éves gyermekek magánhangzóinak akusztikai-fonetikai sajátosságait vizsgáló kutatásának eredményeiről számol be. Korábbi eredmények szerint 5–6 éves óvodások spontán beszédében a rövidhosszú magánhangzók nem különülnek el teljesen minden egyes fonológiai párnál. A 9 éveseknél azonban már szignifikáns eltérés mutatható ki minden magánhangzó-pár esetében. Az életkor előrehaladtával egyre inkább szétválík az egyes magánhangzók formánsszerkezete. Az életkor mind az első, mind a második formánssra szignifikáns hatással van. Az életkor növekedésével egyre rövidebbé válnak a kiejtett magánhangzók is. A 13 éveseknél mért időtartamok már nagymértékben hasonlítanak a felnőttekéire.

A fonológiai tudatosság 4 és 10 éves kor között történő fejlődésével és mérési módszereivel foglalkozik Jordanidisz Ágnes. A fonológiai tudatosság nem más, mint a gyermek metanyelvi tudása. A jelen kötetben ismertetett kísérletben összesen 450 fő vett részt: tipikus fejlődésűek, valamint az írás vagy az olvasás zavarával küzdő tanulók. Az eredmények szerint a magyar gyermekek fonológiai tudatossága egyaránt mutat univerzális és nyelvspecifikus jegyeket is. A fonématudatosság az olvasás megkezdése előtt ugrásszerű fejlődésnek indul. A gyermekek 9–10 éves korra érik el azt a szintet, amikor a három fő terület, a szótag-, rím- és fonématudatosság szintje egyforma lesz. Az itt bemutatott tesztelési módszerrel hozzá lehet járulni a diagnosztikai módszerek sikerességéhez, valamint az elmaradást mutató gyermekek fejlesztéséhez is.

Az utolsó fejezetben Zsák Éva Indira 6, 9, illetve 17 évesek körében vizsgálja a képsorozat alapján való történetmesélés jellegzetességeit. Eredményei szerint a közlések hosszúságában sem a korcsoportok, sem a nemek között nincs szignifikáns különbség. Az életkor előrehaladtával nő a nyelvi sokszínűség, vagyis a szóismétlések aránya kisebb lesz, egyre többfajta határozó jelenik meg, valamint a jelzők használata is gyakoribbá válik. Megfigyelése szerint a hatévesek még nem ismerik fel olyan mértékben a képek közötti ok-okozati összefüggéseket, mint az idősebb korosztály képviselői.

A kötetet ajánlom nemcsak szakmabelieknek, hanem mindazoknak, akik érdeklődnek a gyermeki nyelv jellegzetességei, illetve fejlődési folyamata iránt, valamint azoknak, akik inspirációt szeretnének kapni arról, hogy milyen további kutatási célok tűzhetők ki a gyermeknyelv vizsgálatában.

*Hantó Réka*

**Gyarmathy Dorottya:**  
***Megakadásjelenségek a magyar spontán beszédben***

**MTA Nyelvtudományi Intézet, Budapest, 2017.**

2017 végén jelent meg Gyarmathy Dorottya monográfiája a magyar spontán beszédben előforduló diszharmonias jelenségekről, azaz a megakadásokról. A szerző immár egy évtizede kutatja a megakadásjelenségek gyakoriságát, jellemzőit, előfordulásuk okait a magyar nyelvben. A most megjelent kötet ennek a kitartó kutatómunkának az összegzése, a szerző legfontosabb eredményeinek a bemutatása.

A megakadásjelenségek vagy köznapin nyelven nyelvbottlások vizsgálata több szempontból is fontos eredményekkel szolgálhat. Egyrészt olyan rejtett beszédprodukciós folyamatokra hívja fel a figyelmet, amelyek egyelőre semmilyen más módon nem vizsgálhatók. Másrészt a gyakorlat szempontjából is óriási jelentőséggel bír, hiszen az eredmények több beszéd- és nyelvi

zavar diagnosztizálásához nyújtanak fontos támpontot, árulkodnak az idegnyelvtudás szintjéről, a beszélő egyéni jellemzőiről, és a beszédtechnológiai alkalmazásokban sem hagyhatók figyelmen kívül.

Gyarmathy Dorottya könyvében egyrészt egy általános ismertetést nyújt a megakadások definíciójának kérdéseiről, a típusokba sorolás problémáiról, másrészt részletesen elemez egyes jelenségeket mind a bizonytalanságból adódó jelenségek, mind a hibák közül. Így egyrészt a laikus olvasó is átfogó képet kaphat a témáról, másrészt olyan izgalmas elemzésekbe is bepillantást nyerhetünk, mint például az *izé* előfordulásának jellemzői a magyar beszédben, avagy miről tanúskodnak a perszeverációk.

A kötet első fejezetei a spontán beszéd jellemzőit és a megakadások általános jellemzését tartalmazzák. A szerző bemutatja, hogy milyen beszéd típusok vannak a beszédtervezési folyamatok szempontjából, és foglalkozik a spontaneitás vs. természetesség problematikájával is. Kiemeli, hogy a két fogalom jelentésében nincs átfedés: a spontaneitás a tervezés és kivitelezés szinte egy időben történő voltára utal, míg a természetesség a beszélő számára megszokott beszédet jelenti. A két fogalom különbségének tisztázása rendkívül fontos, amire a szerző kiválóan fel is hívja a figyelmet.

A megakadások természetes velejárói a beszédnek, így univerzális jelenségek, ugyanakkor az egyes nyelvekben más és más beszédtervezési szint okozhat nagyobb nehézséget – így nyelvspecifikusak is. Ennek és a gyakorlati alkalmazások eltérő igényeinek köszönhetően nincs egységes definíciója a megakadásjelenség fogalmának. A szerző ezen probléma bemutatásával kezdi a megakadások tárgyalását, majd bemutatja a kutatásuk rövid tudománytörténeti hátterét, végül hosszabban ír a kategóriába sorolás kérdéseiről. Megemlíti a nemzetközi szakirodalomban megjelenő ellentmondásokat, és ír az eltérő hazai kategorizálások különbözőségéről és hasonlóságairól is. A kötet további részében a hazai szakirodalomban leginkább elfogadott kategóriarendszer szerint csoportosítja a vizsgált jelenségeket. Eszerint két nagy típusba sorolja a megakadásokat: bizonytalanságból adódó jelenségekre és hibákra. A fejezet emellett foglalkozik az önmonitorozással és a hibajavítással is, végül pedig a megakadásjelenségek tudományos hasznáról olvashatunk.

A negyedik fejezet a bizonytalanságból adódó megakadásjelenségek elemzésébe enged a teljesség igénye nélkül bepillantást. Megismerkedhetünk az ismétlések és az újraindítások jellemzőivel, a néma szünet funkcióival, illetve az *izé* és az *úgyhogy* töltelékszói előfordulásaival. Az ismétlések és az újraindítások több szempontból hasonlítanak egymásra (hiszen mindkettőben közös művelet az ismétlés), de más szintű beszédtervezési problémára utalnak. A szerző akusztikai fonetikai mérésekkel elemzi azt, hogy milyen formai jellemzőik vannak ezeknek a megakadásjelenségeknek, illetve a fonetikai megvalósulásuk hogyan utal a funkcióikra. Az összevető elemzés után az újraindítások további, még részletesebb elemzése olvasható. A szerző célja az, hogy kiderítse, milyen okai lehetnek az újraindítások megjelenésének. Az

eredményekből arra következtet, hogy az újraindítások nem minden esetben tekinthetők bizonytalanságból adódó megakadásjelenségeknek, előfordulhatnak a lexikális előhívás problémájából adódó hibajelenségként is.

A megakadások csoportosításakor komoly nehézséget okoz a néma szünetek megítélése. Ezek a jelenségek ugyanis nem minden esetben tekinthetők megakadásnak, hiszen szükség van rájuk a levegővételhez, segíthetik az értelmi tagolást, de retorikai funkciójuk is lehet. Sokszor ugyanazon szünet többféle funkciót is betölthet, ami még inkább megnehezíti a kategorizálást. Gyarmathy Dorottya azon nehéz feladatot tűzi ki célul a néma szünetekkel foglalkozó alfejezetben, hogy akusztikai fonetikai mérésekkel és pozíció szerint elemzésekkel elkülönítse a különböző szünet típusokat. Eredményei szerint igazolható a szünetek időtartamai és a közlésben betöltött funkcióik közötti összefüggés.

A helyettesítő funkciót tekintve talán leguniverzálisabb szavunk az *izé*. Ez a szócska töltelékszónak is tekinthető a terminus pszicholingvisztikai értelmében, mivel nemcsak akkor jelenik meg, ha a beszélőnek nem jut eszébe a kiejteni szándékozott szó, hanem időnyerési céllal is kiejthetjük. Ebben a fejezetben (a szakirodalomban először) a szerző arra keresi a választ, hogy a kétféle funkcióban használt *izék* elkülöníthetők-e akusztikai szempontból is. Emellett egy kérdőíves kutatással felméri az *izé* megítélését is az átlagos beszélők körében.

Az *úgyhogy* kötőszó napjainkban funkcióbővülésen megy keresztül. A kötetben szereplő elemzések szerint a funkcióbővülés jól kimutatható a szó időtartama és a környezetében álló szünetek alapján. Ezzel a szerző egy a jelenben zajló szinkron nyelvi változásra hívja fel a figyelmünket.

Az ötödik fejezetben néhány hiba típusú megakadás, a téves kezdések, téves szótalálások és a perszeveráció jellemzőit ismerhetjük meg. A téves kezdések részletes elemzéséből bepillantást nyerhetünk az önkorrektációs folyamatok működésébe és a téves kezdések feltételezett okaiba. A téves szótalálásokról szóló fejezetben a szerző egy Horváth Viktóriával közös korábbi kutatására alapozva mutatja be, mi történik a mentális lexikon aktiválása során, milyen hibák csúszhatnak a szóelőhívási folyamatokba, illetve hogyan tudjuk korrigálni ezeket.

Az utolsó részletesen elemzett megakadásjelenség a perszeveráció, amely az artikulációs tervezés működési zavara. „Rossz hiba”, mivel egy már korábban kiejtett elem újbóli megjelenését jelenti. A perszeveráció sajátosságainak leírásával ilyen részletesen még nem foglalkoztak korábban, mint a jelen kötetben. A szerző elemzi a perszeveráció típusait, hatótávolságát, az érintett szó szófaját, illetve a javított perszeverációk szerkesztési szakaszát is.

A kötetben bemutatott kutatások igen fontosak és megkerülhetetlenek a nyelvtudomány és a már említett gyakorlati alkalmazások szempontjából. A jó stílusban megírt fejezetek azok számára is érthetővé teszik a problémákat, kérdéseket és az eredményeket, akik egyébként nem járatosak a

pszicholingvisztikai szakirodalomban. Jó szívvel ajánlom tehát a könyv elolvasását nemcsak a nyelvészek és a nyelvvel foglalkozó pszichológusok, hanem mindenki számára, aki érdeklődik az ember csodálatos képessége, a beszédprodukcio működése iránt.

*Bóna Judit*

**Gósy Mária – Krepsz Valéria:  
Morfémák időzítési mintázatai a beszédben**

**MTA Nyelvtudományi Intézet, Budapest, 2017.**

A beszédkorpuszok fejlesztése új korszakot nyitott a beszéd kutatásban. A nagy mennyiségű rögzített beszédet és annak annotációját tartalmazó adatbázisok lehetővé teszik, hogy sok beszélő felvételén történjenek több szempontú, statisztikai eljárásokkal alátámasztott empirikus kutatások. Az MTA Nyelvtudományi Intézet Fonetikai Osztályán fejlesztett BEA és Tini BEA adatbázisoknak, illetve az ELTE Fonetikai Tanszékén létrehozott gyermeknyelvi GABI adatbázisnak köszönhetően a spontán beszéd időviszonyainak sokrétű elemzése is megtörténhetett olyan szempontok alapján, amelyekkel kapcsolatban eddig csak sporadikus kutatások születtek.

Gósy Mária és Krepsz Valéria a morféma időzítési mintázatait elemezte nagy mennyiségű beszédanyagon; a jelen kötet – ennek szintéziseként – az elmúlt években megjelent eredményeket tartalmazza három nagy témakör köré csoportosítva: (i) szavak időzítési jellemzői, (ii) a frázisvégi nyúlás jelensége, (iii) szavak ejtési sajátosságai.

Az első témakör köré négy nagyobb kutatás szerveződött. A szerzők szótövek és toldalékok, valamint igeekötős szavak időviszonyait elemezték. A tanulmányok jelentősége nemcsak abban rejlik, hogy eddig ebben a témában kevés kutatás született a hazai szakirodalomban; hanem abban is, hogy a szerzők elsőként közölnek ilyen jellegű adatokat óvodás gyermekek és tizenéves tanulók beszédéről. Az eredmények szerint a beszélők a szótagszám növekedésével rövidebben ejtették a szótöveket; a toldalékok időtartama azonban állandó maradt. Új eredmény, hogy ez a kiegyenlítődési tendencia az anyanyelv-elsajátításban fokozatosan alakul ki, a felnőttekéhez hasonló mintázat csak a hatéveseknél volt kimutatható.

A második nagyobb rész a frázisvégi nyúlást tárgyalja. A jelenséget a szerzők felolvasásban és nagy mennyiségű spontán beszédanyagon tanulmányozták a szótagszám és a mondatpozíció, a magánhangzó fonológiai időtartama,

a teljes szótag és a beszélő életkora függvényében. Az eredmények szerint szignifikáns különbség volt a felolvasott mondatokban a mondat közepén és legvégén ejtett szavak utolsó magánhangzójának időtartamában; a rövidek a hosszúakhoz képest kevésbé nyúltak meg. A jelenség vizsgálata a spontán narratívákban azt mutatta, hogy a frázisvégi nyúlás a fonológiai hosszúságtól függetlenül bekövetkezett mind a fiatalok, mind az idősek beszédében.

A szavak ejtési sajátosságairól szóló fejezet három kutatás eredményeit ismerteti. A szünet a szóban jelenségét a szerzők a BEA-adatbázis több mint hatvan beszélőjének mintegy 37 órnyi hanganyagán elemezték. A szón belül tartott szünetek részletes leírása ilyen nagy adatmennyiségen szintén új eredmény a hazai szakirodalomban. A gyakorisági elemzések mellett a temporális viszonyok meghatározása is megtörtént a szünet beszédszakaszban elfoglalt pozíciójának és az adott szó szótagszámának függvényében. Az eredmények szerint a szón belüli szünetek időtartama hosszabb volt a beszédszakasz első felében, mint második felében; de az artikuláció leállásának hossza nem független az adott szó szófajától sem.

Gósy Mária tanulmányában három korosztály spontán narratíváiban elemezte a *-ban/-ben* és a *mert* morféma ejtészváltozatait a mássalhangzó-törlődés szempontjából. A /t/ törlődése jellemzőbb volt, mint az /n/ fonémáé. Az eredmények megerősítették azt a szociolingvisztikában korábban leírt tendenciát, hogy a normát/presztízsfórmát legnagyobb mértékben a középkorú beszélők követik – a jelen kutatásban ugyanakkor éppen ez a korcsoport volt a leginkább heterogén a három elemzett morféma ejtésének szempontjából.

A kötet utolsó tanulmányának témája a szavak redukálódásának jelensége a spontán beszédben, amelynek négy típusát tárgyalják a szerzők: (i) lenizációk, vagyis lazítási folyamatok eredményei; (ii) a hiátus feloldása; (iii) egy beszédhang kiesése; (iv) több beszédhang kiesése (szócsonkítás). A szerzők számos nyelvi példát és hangszínképet is közölnek a jelenségek részletes bemutatására.

A tanulmánykötet célkitűzése az volt, hogy bemutassa az elmúlt évek kutatássorozatának eredményeit a szavak időzítési mintázataival kapcsolatban. Az eredmények számos újdonsággal szolgálnak a magyar beszéd temporális viszonyairól, és felhasználhatók az oktatás, a kutatás valamint a beszédtechnológiai alkalmazások területén is.

*Horváth Viktória*